

19 RÉPUBLIQUE FRANÇAISE
INSTITUT NATIONAL
DE LA PROPRIÉTÉ INDUSTRIELLE
PARIS

11 N° de publication :
(à n'utiliser que pour les
commandes de reproduction)

2 771 422

21 N° d'enregistrement national : 97 14669

51 Int Cl⁶ : C 12 N 15/12, C 12 Q 1/68

12

DEMANDE DE BREVET D'INVENTION

A1

22 Date de dépôt : 21.11.97.

30 Priorité :

43 Date de mise à la disposition du public de la
demande : 28.05.99 Bulletin 99/21.

56 Liste des documents cités dans le rapport de
recherche préliminaire : *Se reporter à la fin du
présent fascicule*

60 Références à d'autres documents nationaux
apparentés :

71 Demandeur(s) : CENTRE NATIONAL DE LA
RECHERCHE SCIENTIFIQUE CNRS Etablissement
public à caractère scientifique et technologique — FR.

72 Inventeur(s) :

73 Titulaire(s) :

74 Mandataire(s) : ARMENGAUD AINE.

54 REACTIFS ET METHODES POUR LA DETECTION DE GENES LIES AU COMPLEXE MAJEUR
D'HISTOCOMPATIBILITE D'OISEAUX D'ELEVAGE, TELS QUE LE POULET.

57 L'invention vise des molécules d'acides nucléiques
permettant de détecter ceux des gènes du CMH impliqués
dans les phénomènes de résistance ou de susceptibilité au
développement de tumeurs viro-induites. Les amorces éla-
borées à partir de ces molécules sont utilisables dans une
méthode de génotypage d'oiseaux d'élevage et notamment
du poulet, caractérisée en ce qu'elle comprend
- l'amplification d'un échantillon d'acide nucléique prove-
nant de l'animal à étudier à l'aide d'un ou de plusieurs cou-
ples d'amorces capables de s'hybrider spécifiquement avec
l'acide nucléique d'une région polymorphe des systèmes
Rfp-Y ou B du CMH desdits oiseaux, et
- la détection des produits de PCR obtenus.

FR 2 771 422 - A1

REACTIFS ET METHODES POUR LA DETECTION DE GENES
LIES AU COMPLEXE MAJEUR D'HISTOCOMPATIBILITE D'OISEAUX
D'ELEVAGE, TELS QUE LE POULET

5 L'invention a pour objet la détection de gènes
liés au complexe majeur d'histocompatibilité (CMH)
d'oiseaux d'élevage, tels que le poulet. A ce titre, elle
concerne des molécules d'acides nucléiques permettant de
détecter ceux des gènes du CMH impliqués dans les
10 phénomènes de résistance ou de susceptibilité au
développement de tumeurs viro-induites. L'invention
concerne également les applications de ces molécules
d'acides nucléiques, notamment pour le développement de
tests de génotypage chez les oiseaux d'élevage, en
15 particulier le poulet, et pour la sélection d'animaux
d'intérêt.

Les maladies virales infectieuses sont
redoutées des éleveurs en raison de leur caractère
20 contagieux qui conduit à des pertes importantes
d'animaux.

La vaccination a constitué une prophylaxie
efficace jusqu'à l'émergence de souches hypervirulentes,
25 rendant nécessaire l'identification des haplotypes
résistants.

Diverses méthodes ont ainsi été proposées pour
tenter de sélectionner ceux des animaux qui sont capables
30 de résister à de telles pathologies et ceux qui sont au
contraire susceptibles d'être affectés.

Les techniques les plus utilisées en routine
sont basées sur des polymorphismes sérologiques ou de

type RFLP. Toutefois, ces méthodes ne fournissent pas de connaissances précises sur le phénomène de résistance ou de susceptibilité à la maladie, en particulier par manque de caractère discriminant vis-à-vis des gènes des systèmes B ou Rfp-Y du CMH.

Les travaux des inventeurs sur le séquençage de gènes du CMH a montré la complexité génétique de cette région, ce qui les a conduits à prendre en compte un autre type de polymorphisme, à savoir basé sur la séquence de ces gènes et des régions apparentées, telles que celles de leurs promoteurs et des régions microsatellitaires. Les inventeurs ont ainsi mis au point des moyens pour disposer de molécules oligonucléotidiques hautement spécifiques des polymorphismes observés, permettant d'identifier les parties de gènes, et même les sites impliqués dans le contrôle de la résistance ou de la susceptibilité au développement de tumeurs.

Le caractère spécifique de ces molécules, vis-à-vis d'un gène donné de l'un des systèmes du CMH, en fait des outils discriminants particulièrement fiables pour identifier avec précision la capacité de résistance ou de susceptibilité du poulet étudié, ou d'autres oiseaux, à une infection virale, et pour étudier au niveau moléculaire les séquences du CMH impliquées.

L'invention a donc pour but de fournir des molécules d'acides nucléiques permettant de détecter spécifiquement, chez les oiseaux d'élevage et en particulier chez le poulet, les gènes liés au CMH impliqués dans les phénomènes de résistance ou de susceptibilité au développement de tumeurs viro-induites.

Elle vise également à fournir une méthode et un kit de détection de génotypes de mise en oeuvre aisée en routine.

5 Les molécules d'acides nucléiques de l'invention sont caractérisées en ce qu'il s'agit de molécules, isolées de leur environnement naturel, d'acides nucléiques de gènes codant pour des protéines impliquées dans le contrôle de la résistance ou de la
10 susceptibilité au développement de tumeurs viro-induites chez les oiseaux d'élevage, telles que celles de la maladie de Marek chez le poulet, avec le cas échéant, les régions qui leur sont attachées, telles que celles du promoteur ou microsattelitaires. Le terme gène tel
15 qu'utilisé dans la description et les revendications englobe ces régions.

Ces molécules d'acides nucléiques sont plus spécialement caractérisées en ce qu'elles présentent les
20 séquences d'acides nucléiques de gènes du système B ou du système Rfp-Y du CMH des oiseaux d'élevage, à l'exception des séquences des gènes de classe II B-L, du gène 17.5, du gène 12.3 et du gène B-FIV de classe I, ou sont capables de s'apparier avec l'un des brins d'un gène
25 capable de coder pour une protéine telle que définie ci-dessus dans des conditions faiblement stringentes.

L'appariement dans des conditions de faible stringence auquel il est fait référence ci-dessus est
30 réalisé à température ambiante, dans un milieu 0,1 SSC, avec lavage à température ambiante.

Les gènes de classe II B-L sont décrits dans Immunogenetics 31:179-187, 1990 et Eur. J. Immunol, 1993, 23:1139-1145.

5 Le gène 17.5 appartient à la superfamille des gènes codant pour les lectines et le gène 12.3 à la famille des gènes codant pour des protéines liant la guanine (guanine nucleotide-binding protein). Ce gène est décrit dans Immunogenetics 39:221-229, 1994.

10 Le gène 12.3 est décrit dans P.N.A.S. USA, vol. 86, 4594-4598, juin 1989, Genetics.

15 Le gène B-FIV de classe I est décrit dans Immunogenetics 31:405-409, 1990.

L'invention vise notamment les molécules d'acides nucléiques répondant à ceux des enchaînements de l'un des gènes suivants :

20 . enchaînement du système Rfp-Y

B-FV (figure 1), B-F VI (figure 2) ;

. enchaînement du système B,

25 8.4 génomique (figure 3) ; B-F I (figure 4) ; C121 (figure 5) ; DM (figure 6) ; TAP1 (du début de l'exon 2 à l'extrémité 3') (figure 7) ; et TAP2G (figure 8).

30 L'étude des séquences d'acides nucléiques des molécules définies plus haut a permis de repérer avec précision les blocs de polymorphismes qui doivent être détectés pour établir un génotypage fiable et précis.

En comparant les séquences de ces blocs, provenant de différents gènes d'un même haplotype ou d'un même gène de différents haplotypes, les inventeurs ont pris en considération les enchaînements divergents et élaboré, pour chaque gène, des oligonucléotides complémentaires de ces enchaînements divergents.

On dispose ainsi d'amorces spécifiques et discriminantes vis-à-vis d'un gène donné du système B ou du système Rfp-Y.

L'invention vise tout spécialement les molécules d'oligonucléotides correspondant à ces enchaînements et comprenant une partie de la région polymorphe des systèmes du CMH du poulet ou autres oiseaux d'élevage.

On rappelle que la région polymorphe peut être dans le gène ou dans une région apparentée telle que les régions microsatellitaires ou celle du promoteur.

Selon un mode de réalisation de l'invention, les polymorphismes sont liés à la fonction des systèmes du CMH.

Il s'agit ainsi avantageusement de molécules correspondant à une partie d'un exon. On citera à titre d'exemple des molécules correspondant à l'exon 2 (domaine α 1) des gènes YF du poulet. Un couple d'amorces approprié est constitué par :

Y-F VI α 1 : GGCCCCGGGATGCCGCGGTTTC
Y-F VI α 1, R : ATCCGCTCACCGCCCTGG

Selon un autre mode de réalisation de l'invention, les molécules oligonucléotidiques correspondent à une partie d'une région polymorphe qui n'est pas liée à la fonction des systèmes du CMH. Des régions préférées de ce type sont des microsattellites.

En considérant par exemple, le gène B-FI, des molécules d'oligonucléotides utilisables pour constituer des couples d'amorces correspondent aux enchaînements suivants :

B-FI : 5' CCA GCA GTC ACT GCA CAT AT 3'

B-FI, R : 5' AGG TGG AGT GCG CAA AGT T 3'

Avec les molécules oligonucléotidiques définies ci-dessus et celles élaborées à partir de gènes connus, mais selon la démarche de l'invention, on dispose de jeux d'amorces hautement spécifiques, permettant de déterminer avec précision l'haplotype de l'animal à étudier et de détecter s'il est résistant au développement de tumeurs viro-induites, ou au contraire susceptible d'être affecté.

L'invention vise donc également une méthode de génotypage d'oiseaux d'élevage et notamment du poulet.

Cette méthode est caractérisée en ce qu'elle comprend

- l'amplification d'un échantillon d'acide nucléique provenant de l'animal à étudier à l'aide d'un ou de plusieurs couples d'amorces capables de s'hybrider spécifiquement avec l'acide nucléique d'une région polymorphe des systèmes Rfp-Y ou B du CMH desdits oiseaux,

et

- la détection des produits de PCR obtenus.

Une simple comparaison des résultats obtenus avec un référentiel établi au préalable permet de
5 déterminer rapidement l'haplotype de l'animal.

L'échantillon d'acide nucléique est constitué en particulier par de l'ADN génomique extrait de matériel biologique de l'animal à étudier ou par ce matériel même, en particulier par du sang de l'animal. Il peut s'agir en
10 variante d'ADNc, d'ARN ou encore de PNA (polypeptides nucleic acids).

Les amorces sont élaborées à partir des molécules oligonucléotidiques définies ci-dessus et,
15 d'une manière générale, de tout gène (et région apparentée) codant pour une protéine impliquée dans le contrôle de la résistance ou de la susceptibilité aux tumeurs viro-induites chez les oiseaux d'élevage et notamment de poulet, en particulier les gènes B-L de
20 classe II, 17.5, 12.3 et B-FIV de classe I.

Il s'agit par exemple d'amorces de régions microsatellitaires permettant de détecter des haplotypes du complexe B, telles que celles élaborées à partir du gène B-FI, et évoquées ci-dessus, ou d'amorces permettant
25 de détecter des haplotypes du système RFP-Y, et élaborées à partir du gène 17.5, comme le couple :

17.52 : CAG GAT CTG CAC TGG CCA ATA

30 17.5, R1 : GAA TGG CGG TGC TTC CGT GCC TGG

La détection des produits de PCR est effectuée selon les techniques classiques. Ces techniques

comprennent le séquençage, l'électrophorèse, les hybridations avec analyse SSOP ou SSCP.

5 Cette technique sera avantageusement choisie selon la nature du polymorphisme impliqué. Ainsi, dans le cas de polymorphisme de type microsatellite, on détectera avec avantage les produits de PCR selon leur taille en ayant recours aux techniques d'électrophorèse.

10 Lorsque le polymorphisme ne concerne que quelques nucléotides, voire un seul nucléotide, on aura plus spécialement recours, aux fins de différenciation des haplotypes de produits de PCR, aux techniques d'hybridation (analyse sur membrane à l'aide de sondes
15 spécifiques des séquences d'haplotypes, SSOP ou Sequence Specific Oligonucleotide Probe), de migration différentielle des échantillons dénaturés (SSCP ou Single Strand Conformational Polymorphism), ou de séquençage. De manière générale, cette dernière technique est préférée
20 compte tenu de la simplicité de sa réalisation.

L'invention fournit ainsi une technique simple et rapide d'établissement du profil génétique d'un grand nombre d'animaux à étudier, ce qui permet de déterminer
25 les haplotypes et de sélectionner ceux d'intérêt en vue d'un élevage.

De plus, chaque type de gène pouvant être discriminé en utilisant des amorces présentant la
30 spécificité requise et son appartenance au système B ou Rfp-Y pouvant être établie, il est possible d'effectuer des études fondamentales plus complètes.

L'invention vise également un coffret ou trousse pour détecter le génotype du poulet ou autre oiseau d'élevage selon la méthode définie ci-dessus.

5 Ces coffrets ou trousse sont caractérisés en ce qu'ils comportent les réactifs nécessaires pour la réalisation d'au moins une PCR et du test de révélation.

10 En particulier, ils comportent les amorces pour la PCR, un témoin positif de la réaction, ainsi qu'une notice d'utilisation.

15 Les amorces se présentent sous forme lyophilisée ou en solution ou, selon le mode de détection, sur un support. Le support peut être, de manière classique, une plaque multipuits ou se présenter sous forme de puces à ADN.

20 D'autres caractéristiques et avantages de l'invention sont exposés dans les exemples qui suivent, dans lesquels il est fait référence à la figure 9 représentant une photo d'électrophorèse de produits de PCR illustrant le test de génotypage de l'invention. On rappelle que les figures 1 à 8, déjà évoquées ci-dessus,
25 illustrent les séquences de gènes selon l'invention.

Exemple :

30 Etude d'haplotypes Rfp-Y du poulet à l'aide d'amorces microsatellitaires.

- amplification avec le Kit Expand™ High Fidelity PCR System

. Avec les amorces 17.5 R1/17.52

ADN génomique : 1 μ g

Oligos prendre : 0,3 μ M

5 dNTP : 8 μ l

qsp H₂O 50 μ l

On ajoute 50 μ l de Mix 2 en mélangeant.

Mix 2 : 0,75 μ l d'enzyme

10 10 μ l TP10X avec MgCl₂

qsp H₂O 50 μ l

Programme d'amplification :

15 30 Cycles

94°C	94°C	65°C	72°C	4°C
2'	30''	1'	1'	∞

20 . Avec B-FI/B-FI, R :

ADN génomique : 1 μ g

Oligos prendre : 0,3 μ M

dNTP : 8 μ l

25 qsp H₂O 50 μ l

et ajouter 50 μ l de Mix 2 en mélangeant.

Programme d'amplification :

30 30 Cycles

94°C	94°C	60°C	72°C	4°C
2'	30''	1'	1'	∞

- révélation par électrophorèse sur gel d'agarose ou par séquençage.

5 Le test a été appliqué à 9 haplotypes de poulet, sélectionnés sérologiquement pour le complexe B. Il s'agit des haplotypes B4, B5, B7, B12, B13, B14, B15, B21 et d'un haplotype inconnu BX.

10 Plusieurs individus d'un même type ont été étudiés pour B12 (6 individus), B13 (3 individus), B14 (4 individus), B21 (4 individus) et un seul individu pour les autres haplotypes.

15 La figure 9 donne une photo d'électrophorèse sur gel d'agarose à 1 % des produits de PCR obtenus à l'issue de l'étape d'amplification.

20 Les pistes 1 et 27 correspondent aux marqueurs de taille et les pistes (2 à 25) aux produits de PCR des haplotypes suivants : piste 2 : B4 ; piste 4 : B5 ; piste 5 : B7 ; pistes 6 à 11 : B12 ; pistes 12, 13, 14 : B13 ; pistes 15, 16, 17, 18 : B14 ; piste 19 : B15 ; pistes 20, 21, 23, 24 : B21 ; piste 25 : BX (absence de détection pour les pistes 3 et 22).

25 L'examen de cette figure montre que les individus qui ont l'haplotype B12 donnent une même bande et sont donc bien homogènes. La même observation s'applique aux individus B14. En revanche, avec B21, on constate que les profils sont différents, ce qui démontre l'inefficacité de l'approche sérologique. Compte-tenu de
30 la position de la bande de BX, on détermine qu'il s'agit d'un haplotype B4.

L'application pratique de cette méthode revient à soumettre les individus naturellement résistants au protocole décrit ci-dessus en prenant en compte les deux systèmes Rfp-Y et B du CMH et à ne sélectionner parmi des animaux à tester que ceux dont le profil correspond à celui des animaux résistants.

L'invention fournit ainsi les moyens de vérifier l'homogénéité des animaux et d'effectuer des sélections rigoureuses en prenant en compte chaque système du CMH, et dans ces systèmes les gènes recherchés.

REVENDEICATIONS

1/ Molécules d'acides nucléiques isolées de leur environnement naturel, de gènes codant pour des protéines impliquées dans le contrôle de la résistance ou de la susceptibilité au développement de tumeurs chez le poulet, telles que celles de la maladie de Marek, et de régions apparentées auxdits gènes caractérisées en ce qu'elles présentent les séquences d'acides nucléiques de gènes du système B ou du système Rfp-Y, correspondant au complexe majeur d'histocompatibilité des oiseaux d'élevage à l'exception des séquences des gènes de classe II B-L, du gène 17.5, du gène 12.3 et du gène B-FIV de classe I, ou sont capables de s'apparier avec l'un des brins d'un gène capable de coder pour une protéine telle que définie ci-dessus dans des conditions faiblement stringentes.

2/ Molécules d'acides nucléiques selon la revendication 1, caractérisées en ce qu'elles répondent à l'un des enchaînements suivants :

. enchaînement du système Rfp-Y

B-FV (figure 1), B-FVI (figure 2) ;

. enchaînement du système B,

8.4 génomique (figure 3) ; B-FI (figure 4) ; C121 (figure 5), DM (figure 6), TAP1 (du début de l'exon 2 à l'extrémité 3') (figure 7), et TAP2G (figure 8).

3/ Molécules d'acides nucléiques selon la revendication 1 ou 2, caractérisées en ce qu'elles correspondent à une partie des séquences définies dans les revendications 1 ou 2, cette partie étant spécifique

et discriminante pour un gène donné des systèmes B et Rfp-Y.

5 4/ Molécules d'acides nucléiques selon la revendication 3, caractérisées en ce qu'il s'agit de molécules d'oligonucléotides correspondant à une partie de région polymorphe des systèmes du complexe majeur d'histocompatibilité du poulet.

10 5/ Molécules d'acides nucléiques selon la revendication 4, caractérisées en ce qu'il s'agit de molécules d'oligonucléotides correspondant à une partie d'exon.

15 6/ Molécules d'acides nucléiques selon la revendication 4, caractérisées en ce qu'il s'agit de molécules d'oligonucléotides correspondant à une partie de région polymorphe qui n'est pas liée à la fonction des systèmes du CMH, telle que les régions
20 microsatellitaires.

7/ Méthode de génotypage d'oiseaux d'élevage et notamment du poulet, caractérisée en ce qu'elle comprend

25 - l'amplification d'un échantillon d'acide nucléique provenant de l'animal à étudier à l'aide d'un ou de plusieurs couples d'amorces capables de s'hybrider spécifiquement avec l'acide nucléique d'une région polymorphe des systèmes Rfp-Y ou B du CMH desdits
30 oiseaux,

et

- la détection des produits de PCR obtenus.

8/ Méthode selon la revendication 7, caractérisée en ce que les amorces sont élaborées à partir des molécules selon l'une quelconque des revendications 3 à 6, et de tout gène (et région apparentée) codant pour une protéine impliquée dans le contrôle de la résistance ou de la susceptibilité aux tumeurs viro-induites chez les oiseaux d'élevage et notamment de poulet, particulièrement les gènes de classe II B-L, 17.5, 12.3 et B-FIV.

9/ Méthode selon la revendication 7 ou 8, caractérisée en ce que la détection des produits de PCR est effectuée par séquençage.

10/ Coffret ou trousse pour le génotypage d'oiseaux d'élevages et notamment du poulet, caractérisé en ce qu'ils comportent les réactifs nécessaires pour la réalisation d'au moins une PCR et du test de révélation, selon la méthode de la revendication 8 ou 9, en particulier les amorces élaborées à partir des molécules d'acides nucléiques selon l'une quelconque des revendications 3 à 6.

Figure 1

BF V

B-F V

GGC CCC GGG ATG CCG CGG TTC GTG ATC GTC GGG TAC GTG GAC GAC AAA ATC TTC GGT
 ACC TAC AAC AGT AAG AGC AGG ACT GCA CAG CCT ATC GTG GAG ATG CTG CCG CAG GAG
 GAC CAG GAG CAC TGG GAC ACG CAG ACC CAG AAG GCG CAG GGC GGT GAG CGG GAT TTT
 GAC TGG AAC CTG AAC AGG CTG CCG GAA CGC TAC AAC AAA AGT AAA GGT GAG CGT GGG
 GGA AGC TGC AGC GCG ATG CGT CTG GGA CAG GAG CTC TGT GTG CCG AGG GTG TCC GCC
 AGC CCC ACT GAG GTG TGG CCG TGC CCC ACG CCC AGC TGT GCT GGG CCG TCC ATG TGT
~~GCT~~ GGC ACT GTC CCT GGG CCG CCC TGC TCC TGC GCC CAC CCA CCC CAC CCC AGC CTC
 ATG GCA CTC GCG GTG CCC CAC AGC CCT AGA AGC CTC TCA CCT ATT ACT CTG GCT GTG
 CCT CAG GGT CTC ACA CGA TGC AGA TGA TGT TTG GCT GTG ACA TCC TTG AGG ACG GCA
 GCA TCC GAG GGT ACG ATC AGT ATG CAT TTG ATG GGA GGG ACT TCC TTG CCT TTG ATA
 TGG ACA CGA TGA CGT TCA CCG CGG CGG ATC CAG TGG CTG AAA TCA CCA AGA GGA GAT
 GGG AGA CAG AAG GGA CGT ATG CTG AGA GAT GGA AGC ATG AGC TGG GGA CTG TCT GTG
 TTC AGA ACT TGA GGA GAT ACC TGG AGC ATG GGA AGG CAG CGC TGA AAA GGA GAG GTG
 AGG ATG GGA GGG GGA CGT GGG GCT GGG CTG GGT GTG GGG CAG AGG CTC AGT GTG GGG
 TGC TCA GCC CGG CCC ACA ACG TCA CCC ACC TGC AGT GCA GCC CGA GGT GCG AGT GTG
 GGG GAA GGA GGC CGA TGG GAT CCT GAC CTT GTC CTG CCA CGC TCA CGG CTT CTA CCC
 GCG GCC CAT CAC CAT CAG CTG GAT GAA GGA CGG CAT GGT CCG GGA CCA GGA GAC CCG
 CTG GGG GGG CAT CGT GCC CAA CAG CGA TGG CAC CTA CCA CGC CTC GGC TGC CAT TGA
 TGT GCT GCC GGA GGA TGG GGA CAA GTA TTG GTG CCG CGT GGA GCA CGC CAG CCT GCC
 CCA GCC TGG TCT CTT CTC ATG GGG TGA GCT GGC AGC GTG GGG CAC GTG GGG TTG GGA
 TTC GCA GGC TGC CCC TTC CTT TAC TGA CAA CGG CGC TCT CCT CCA GAG CCG CAG CCC
 AAC CTG ATT CCC ATT GTG GCA GGG GCG GTC GTT GCC ATC GTG GCT GTC ATC GCT GCG
 GTC GTT GGA TT

B-FVI

GGC CCC GGG ATG CCG CGG TTC GTG ATC GTC GGG TAC GTG GAC GAC AAA ATC TTC GGT
ATC TAC GAC AGT AAG AGC AGG ACT GCA CAG CCC ATC GTG GAG ATG CTG CCG CAG GAG
GAC CAG GAG CAC TGG GAC GCG CAG ACC CAG AAG GCC CAG GGC GGT GAG CGG GAT TTT
GAC TGG TTC CTG AGC AGG CTG CCG GAA CGC TAC AAC AAA AGT GGA GGT GAG TGT GGG
GGA AGC TGC AGC GCG ATG CGT CTG GGA CAG GAG CTC TGT GTG CCG AGG GTG TCC GCC
AGC CCC ACT GAG GTG TGG CCA TGC CCC ACG CCC AGC TGT GCT GGG CCG TCC ATG TGT
GGT GGC ACT GTC TCT GGG CTG CCC TGC TCC TGC GCC CAC CCA CCC CAC CCC AGC CTC
ATG GCA CTC GCG GTG CCC CAC AGC CCA AGA AGC CTC TCA CCT ATC ACT CTG ACT GTG
CCT CAG GGT CTC ACA CGA TGC AGA TGA TGA TCG GCT GTG ACA TCC TGG AGG ACG GCA
GCA TCC GAG GGT ACG ATC AGT ATG CAT TTG ATG GGA GGG ACT TCC TTG CCT TTG ATA
TGG ACA CGA TGA CGT TCA CCG CGG CGG ATC CAG TGG CAG AAA TCA CCA AGA GGA GAT
GGG AGA CAG AAG GGA CGT ATG CTG AGA GAT GGA AGC ATG AGC TGG GGA CTG TCT GCG
TTC AGA ACT TGA GGA GAT ACC TGG AGC ATG GGA AGG CGG CAG TGA AAA GGA GAG GTG
AGA ATG GGA GGG AGA CGT GGG GCT GGG CTG GGT GTG GGG CAG GGG CTC AGT GTG GGG
TGC TCA GCC CGG CCC ACA ACA TCA ACC ACC TGC AGT GCA GCC CGA GGT GCG AGT GTG
GGG GAA GGA GGC CGA TGG GAT CCT GAC CTT GTC CTG CCA CGC TCA CGG CTT CTA CCC
CCG GCG CAT CGC CAT CAG CTG GAT GAA GGA CAG CAT GGT CCA GGA CCA GGA GAC CCG
CTG GGG GGG CAT CGT GCC CAA TAG GGA TGG CAC TTA CCA CAC TTC GGC TGC CAT TGA
TGT GCT GCC GGA GGA TAG GGA CAA GTA TCG GTG CCG CGT GGA GCA CGC CAG CCT GCC
CCA GCC TGG CCT CTT CTC TTG GGG TAA GCC TGG CAG CGT GGG ATG TGT GGA GTT GGG
ATT TGG GGG CCG CCC CTT TGT TTA CTG ACA ACG GTG CTC TCC CCC AGA GCC GCA GCC
CAA CCT GAT CCC CAT TGA GGC TTG GCT GGT CGT CCC CTT GGT GGT TCT CTT CGT TGC
TTT GAT TGC ATT

GGA TCC GGG GTG GGT GGC AGT GGC TGT GTT TAG GTC GGC CTG TGG GGA AAG
CCG GGT TGT CCC ACC CAT GTC CCC TCT TCC AAC ACT GTT CCT GAA TGA GTT
TTC CCT CTC CGA CCC TTT TTT TAA TGG GTT TCA GGG ATT TAA AAT TAA TAT
TGA CGA AGT GAC GGA GGG GGT GGG GCC ACA GCG GAG CCG AAA GCG AAA GCA
GCG GAG AGC AAT GGC TGC GGG GCT GCG GCT GCT GCT GGC GGG TGA GAC CCG
ACC CCC CCC GGC CCC CTC ATG TCC CAC CAC CCA TAT CGC CCC CCC CCC TCC
TCC TCG CCC CAT GCT GAG CCT CTC CCC CAC CCC CAG GGC TCT GCT GGT CCC
AAT TTA GGG TGG AAG ACG CCG CCT CCC CTC CGC CCC CCC CCG CTC CGG TGC
GCT GCG CGC TGC TGG AGG GGG TGG GGC GCG GGG GAG GGC TGC CGG GGG GGG
GCA ATG CCC GTC CTG CAC TGC TGC GCT TTG GGG GGG ACG CGG AGA CCC CTC
CCG AAC CCG GCC CGG AGC CCG AAG TCA CCT TCA ATG TCA GCG GTA CGT GGG
GAC CCC CGT CAC TGT GCT GTG CGC CTC CTT TAT CCC CAC CCC CCT CCA TGT
CCC CAT CTC CTT TAC TTC CCA CAA TGC TCC CAT CCC CCC CAG AAT GTC CCC
AGA GTC CCC CAA ACC CCC ATG ACC CCC CCC ACG ACC CCT GGT TCC CAT TAC
CCT CTC ACG TCC CCC AGT GTC CCC AAG ATT CCC ATT ACT CCC CGT ATC CCC
ATT ATC CCC AAA ATG TCC CCC AAT GTT CCC ATC ACC CCA ATG TTC CCA AGG
TCC CTA TCG CTC CTC AAT GTC GCT ATG ATC CCT ATT CCC AAA ATG TCA CCA
ATG TCC CCA AAA TCC CCA TTA TCT CCC ACC TCT CCA AAG TCC CCA AGA TCC
CCA TTA CCC CCA ATA TCC TCA TTA CAC CCC AAA TGT CCC CAA TGT CCC CTC
CAT GTC CCC CAG AGA CCC CAT TAG CCC CAA TAG CTC CCA AAC TGT CCC CAG
TGT CCC CAT TAA CCC CAA AAT GAC CCC ATT ACG CCC CAC ACC CCT CCC AAC
CCC ATG CCC TCA GAC CCC TTC ATC CCT CTC ACT CCT CTC TCC CTC GCA GAC
CCC TGG GGG ACT CTA GCC CCA CTC GGG TCC CCC CCC GGA CTC CCC CCA GCT
GCG AAC TGA ACC CCA CGA ACC CCC AGA CCG GCT CTG ACC CAT GGA GCC GCC
CTC TGC ACC CCG ACG CCC GCA GCC CCC CAA CCG CGG GGG GGC AGT GGT GGG
TGG CGG CGG TGG GGA CCC CGC AGT ACG GTG TCA CTG CGC TGC TGC AGG GGG
GGA TGG GCA CAG AAG GAA CCA TCA CTG CCG CCG GTA AGG GGG AAC TTG GGG
TGT CCC TCC CTG GGT GTC CCC ATG TCC CTA TCT GTC CCC CAG TGT GTC CCC
ATT TGT CCC CTC CTC TGC ATG TGT CCC AAT GTC TCC ATA CAT CCC ATA ATA
ACC ATA TGT CCC CAC TCA TCC CCA TAT TCC CCA TGT GTC CCC ATA TCC CCA
CAC ATC CCA GTG TGC CCC AAC ACA TCC CCA TGT GCC CCC CCC CAT GCA TCA
CTA CCA TCC CCC TAT CCC CCA AGT GTC CCT GTG TCC CTG CAG TTT CTC CCT
GTC CTC ATG TGT TCC CAT GTC TCC ATG TCA CTG TGT CCC CGT GTC CCC ACA
CAT CAC CAT GCC CCC CAC TGC AGC GCC CCC ATG TCC CTT CAC CTC TCC ATG
TCC CCC AGT GTC CCC TAT CCC CTC ATT GTC CCC ATG CCC CCT CAC CTC CCC
GTG TCC CCC GTG TCC CTA TGT TCC CCT GGT GTT TCC ATG TCC CCT CAT GCC
CCC ATG TCC CCT CAT GTC CCC ATA TCC CCC AGT GTC CCC ATG TCC CTT CAC
CTC CCC ATG TCC CCC AAT ATT CCC ATA TCC CCT CAC CTG CCC ATT TCC CCC
CGA TGT TCC CAT GTC CCC GCA CCT CCC CAT GTC TTC ACA GTG GCC CTG GCG
GTG CTC ACC CAC ACC CCG ACC CTC CGG GCC CGT GTG GGG TCC CCC ATC CAC
CTG CAC TGC GCC TTC GCT GCC CCC CCA TCC TCC TTT GTC CTC GAG TGG CGT
CAC CAG AAC AGG GGT GCG GGG AGG GTC CTG CTG GCC TAT GAC AGT TCC ACC
GCC CGC GCC CCC CGC GCC CAC CCC GGG GCC GAA CTG CTG CTG GGG ACA CGG
GAT GGG GAC GGG GTG ACA GCG GTG ACA CTG CGG CTG GCG CGG CCA TCA CCG
GGG GAT GAG GGC ACC TAC ATC TGC TCC GTG TTC CTG CCC CAC GGG CAC ACA
CAG ACA GTG CTG CAG CTC CAC GTC TTT GGT GCG TCC ATG TGG GGC AGG CGG
TGT TCC TAT GGG GTG TGG GGT TGG GCA GTG TTC CTA CGG AGT GTG TAT GAC
TGG GTG GTA TTC CTA TTG GTC AGA TAG GAC ATA TGG GAG CAG GCG GTA TTC
CTA TGG GGC TGT AGG GTG GAT GGG ACT GGG TGA TAT TCC TGT GGG GGC TGT

AGG GTG GAT GGG ACT GGG TGG TAT TCC TAT GGA GGC TAT AGG GTG GAT GGG
ACC GGG TGG TAT TCC TAT GAG GAC TAT AGG ATG GGG TGG CAT CAT CCC ATA
GTT CAC CTG TAG GTT TAT AGG GGG GGA TGA GCC CTA TAC AGC GTA TGG GCT
ATA TGG ACC GAT GTC CCC CCA CAT GTC TCC AGA GCC CCC CAA GGT GAC GCT
GTC CCC GAA GAA CCTGGT GGT GGC CCC GGG GAC GTC AGC AGA GCT ACG CTG
CCA GTC TGG CTT CTA CCC CTT GGA TGT GAC GGT GAC GTG GCA GCG CCG CGC
CGG GGG CTC GGG GAC ATC ACA GTC ACC CAG GGA CAC AGT GAT GGA CAG CTG
GAC TTC AGG TCA CCG CCA GGC AGC CGA TGG AAC CTA CAG CCG GAC GGC GGC
AGC ACG GCT GAT CCC CGC ACG CCC CCA ACA CCA CGG GGA CAT CTA CAG CTG
CGT TGT CAC CCA CAC TGC ACT GGC CAA ACC AAT GCG TGT CTC CGT CCG ACT
GCT CCT GGC TGG TGA GGG GGG ATG TGG GGA TAT TGG AAA CAC GTG GAG GTA
TTG GGA TGC TGG GAC CAT GGT TAG GAG GGT CTG AGG GAC ATC AGG ACC ATG
GCC TGG GAC AAT GGG AGA TCA TGG ATT TGG GTT GGG GAC CCC ACC CAG GAT
GGT GAC ACT GTG CTT AGG GCT GTC GTT GTC CCC ACA GGC ACC GAG GGA CCG
CAC CTG GAG GAC ATC ACG GGG CTC TTC TTG GTG GCC TTT GTC CTC TGT GGC
CTC ATC CGT TGG CTC TAC CCT AAA GGT GAG TGC TGT TCC CAC ATC CCA GTG
CCC CCA CAT CCT CAC ACC CCA ATA TCC CAA TGG CCC ATG TCC CCA TGA GCA
ATG TCA CTA TGT CCC AAT ATC CTA ATG ATG CTG TGT ACC CAT GTG TCC CCA
TGT CCC TAT TCC ACT CAC TCT TTC TCT CCC CTC AGC TGC ACG ACC CAA AGA
GGA AAC CAA GGT AAC ATT CCT CCC CAA AAA CCC CAA ATC CCC CAA AAC ACC
TCC AAG CAC CCC AAA ACT CAC CAT TCT CAT TCC CCC CCC CCC CCC CCC
CCC CAT GCC TTG CAG AAA TCG CAG TGA CCT CCA CTC CAG CTC TCA GCA CCT
CAG CTC CAG ATA AAG AGT TTT TCA CCC CAA AGT TAT ATA TGT GTG GTG-GTG
TCC CCA CAG ATC TGG GTG CAG AGG GGG GAG AAA TGG GGG CAA ACT GGG AGC
AGT GGG AGC AGT GGG AGG AAG TCC TGG GTT GGT GAG GCA GAT GAG TGG CAC
CTG GGG ACA TCT GGG TGC CAT CCC TTG TGG ACA TCT GGG TGA CAC TGC ATT
GCC TTG GGT GAC ATT GGG ATC CTC AGG TCA CTG CAG

GT CGA CGGGAT CTGGATAGGT CGT CAGT CAT CCTAATTAAGGAGGGA CAA CAGTGAATGGG
GAGGAG CCGATGA CT CAGGCTGGGAGTGGTGAAT CCCAGAGGTTT CCT CTG CTGT CAGTGA C
T CCGTG CTTT CG CTTT CG CTT CA CAA CCTC GGGAG CG CATT CTG CTTGG CG CCCGATGA C
GT CA CATAAA CCCCCGA CTG CCATTGG CGGAGAGG CGA CGGAGGAG CCAATGGGGG CG CGG
GG CGGGG CGGAGGAGTAGGAAAAG CTGAAGGA CGTG CG CTGGGTG CGG CGGACTTGAGAGT
G CAG CGGTGTGAGG CGATGGGG CCGTG CGGGG CG CTGGG CCTGGGG CTG CTG CT CG CCG CC
GTGTG CGGGG CGG CGG CCGGTGAGTG CGG CCGGA CCGGGA CCCCT CCCCG CCTGTAA CCCC
A CCCC GGG CTGTG CCGTGGGAT CCT CAGAC CCCCCA CCG CGG CT CA CGG CCT CG CTG CCG
T CCG CCCCCG CAGAG CT CCATT CCCTG CGGTA CGT CCATA CGG CGATGA CGGAT CCGG CC
CCGGG CTG CCGTGGTT CGTGGA CGTGGGGTA CGTGGA CGGGGAACT CTT CGTG CA CTA CAA
CAG CA CCG CG CGGAGGTA CGTG CCG CG CA CCGAGTGGATGG CGG CCAA CA CGGACCAG CAG
TA CTGGGATGGA CAGA CG CAGAT CGGA CAGGG CAATGAG CGGAGTGTGGAAGTGAG CT TGA
A CA CA CTG CAGGAA CGATA CAA CCAGA CCGG CCGT GAG CA CGG CCGGGG CCG CGG CT CCGT
GGGTGTGGGATGGG CT CCATGG CG CAGTG CCG CCA CA CCCCC CAGG CCTGG CCCTG CCGG
G CGG CA CCGT CCGGGG CTG CCGT CA CAG CCCCCA CCG CG CT CGGGGTG CCG CGT CCGGG
GGGA CCCCCA CCG CCG CTG CAGTGGGAG CCGGAG CCGGAGGGG CCCCT CA CCCCCT
G CCGG CTGTGTTT CAGGGT CT CAAA CCGTG CAG CTGATGTA CGG CTGTGA CAT CCT CGAG
GATGG CACCAT CCGGGGGTAT CAT CAGACAG CCTA CGATGGGAGAGA CTT CATTG CCTT CG
A CAAAGG CA CGATGA CGTT CA CTG CGG CAGTT CCAGAGG CAGGTT CCA CCAAGAGGAAAT
GGGAGGAAGGAGGTGTTG CTGAGAGGTGGAAGAGTTA CCTGGAGGAAA CCTG CGTGGAGGG
G CTG CGGAGATA TGTGGAATA CGGGAAGG CTGAG CTGGG CAGGAGAGGTGAG CGGGGT CGG
GGTGGGGGGGGGGGGGGG CGGA CG CAGTGTGGGGT CGGA CGTGGGG CGGGGG CT CAT CGTG
GGGAG CT CAG CCGG CCCT CA CTG CCG CCA CCA CAGAG CGG CCTGAGGTG CGAGTGTGG
GGGAAGGAGG CTGA CGGGAT CCTGA CCTTGT CTG CCG CG CT CA CGG CTT CTA CCG CGG C
CCAT CG CCGT CAG CTGG CTGAAGGA CGG CG CGTG CGGG CCAGGA CG CCGAGT CGGGGG
CAT CTGTG CCAA CGG CGA CGG CA CCTA CCA CA CCTGGGT CA CCA CTGA TG CG CAG CCGGG
GA CGGGGA CAAGTA CCAGTG CCG CGTGGAG CA CG CAG CCTG CCCCAG CCGG CCT CTA CT
CGTGGGTGAGTGAGGGGATGTGGGG CTGGGGG CTG CGGG CTG CCCCTT CCCCTG CTGAT
GG CCGG CT CT CCCCAGAG CCG CCA CAG CCA CCA CCTGGTG CCGAT CTGG CGGGGTGG C
CGT CG CCATTGTGG CCA CT CG CCA CTGTGGTGGTGTGGATT CAT CAT CTA CAGA CG CCA C
G CAGGTAAAAG CAGAGGGGTG CAGG CGGG CAGTGGGG CTGTAGGGGGAT CTGGGT CCCCC
CTTGGGAG CCCCCA CCTGG CTGTGATGTGAA CCTGTGATGAAG CAT CT CT CTGT CTG CAG
GGAAGAAGGGGAAGGG CTA CAA CAT CG CG CCGGTGAGTGATGAGGG CAG CG CTGT CCCCC
A CCT CTG CCGAGTG CAGGGTGGT CCTGGGGT CCCTG CTTT CT CCAAGGTA CCGATT CCT
GGTG CTTGGGG CTG CT CCATG CCGCATAGGGAG CA CAGGG CTGGAT CT CA CAG CTGTT CCT
CCCTTATAGA CAGGGAAGGTGGAT CCAG CAG CT CGAG CA CAGGTG CCGTGTGGGG CTGTGG
GTTGGGAGGGGT CCGTGTG CT CT CTGTGGTACTG CCGAGGG CTGGG CTATG CTGGGG CT CT
G CGGGGAGAC CCGGAG CAGAGGGT TGGGATGTGAAC CTGG CCGGCTGGGA CAT CAT CCC
TT CT CAT CCA CAGGGAG CAA CCGG CCA CTGAGTG CTGTG CTT CAG CCTG CAAGGAG CC
AA CAGT CCA CA CAG CATT TGGGGT CCGTGATGGA CA CAG CCGCAT CCT CCTGACCT CT CA
CAT CT CATT CTG CTT CCTATG CTGACTGTTATG CTTTG CCTG CA CTG CTT CCTGTGAAATA
AAATGATGGG CCA TT CTGTG CT CAG CTTG CCTG CATT CTG CA CAGTG CTGTGGTTGGGGAT
GGGGTGGGTGAGAGGA CCGTGT CCAGTTTGG CTG CT CAGGGTG CAGATGTGG CCCTGTG CT
GAGTA CCA CTG CCCT CCCCCCTAT CTG CCTG CTG CT CA CT CCCCCCT CCTGTAC CCCCCAT
CCCTT CT CACCT CT CCT CTGTGAA CCGCATG CTGGTGGTTGCTTGCT CCCTGT CCTGG CAG
AA CT CT CATT TT CCAATGG CAT CCCTGGGTGTTGGGATGTGGT CT CCTTGGT CCT CCCCC
CAG CAGT CA CTG CA CATAT CCCCCCA CTT CCCCCCTAGGTTGTTGT CCA CAG CA CT CCT
ATTT CCGT CT CCCCCCCCCCCCCCCCCCCCCCG CCG CCA CAG CTG CCT CTG CAAT CCT CA C
CCTTG CCA CCA CAA CT TTG CG CA CT CCA CCT CCCT CAT CCG CCCTT CCCCCAG CT CT C
CTGT CCCTG CTGG CCCCCCT CCCCCCCCCCATTTGTAC CCA CCA CCAATAAATATGTTT
GTT CTG CTGCCCT CCAG CCGT CT CCTGGTTTATTT CCCCCGATTTGTTGTTGGGG CG

6/31

2771422

TCCGCTCTTCAACCCTGGGGGAAGGGGCTCTGGGGGTCCCTCATTCTCCCTGCACTTCTTA
CAGCACCGGGACTCCCGCGCTGAGATCCCATCACACCGGGTACAAACATGCGGCTTTATT
CCCAGTTCTGTGTCCCACCCCGGCCCTGCTGGCACTCAGTGGCACCGCAGTCCATGCAGT
GGCCJTTGTGTGTGTGTGTAAGCAGCGGTAC

ATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGC
ACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCCATAGTTGCCTGCAAC
TCCCCGTCGTGTAGATAACTACG.TACGGGAGGGCTTACCATCTGGCCCC
AGTGCTGCAATGATACCGCGAAGACCCACGCTCACCGGCTCCAGATTTAT
CAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCA
ACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGT
AAGTAGTTCGCCAGTTAATAGTTTTCGCAACGTTGTTGCCATTGCTGCAG
GCATCGTGGTGTACGCTCGTCTGTTTGGTATGGCTTCATTACGCTCCGGT
TCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGC
GGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAG
TGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATG
CCATCCGTAAGATGCTTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATT
CTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAACAC
GGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGA
AAACGTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTTGAGATC
CAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTA
CTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCCAAAATGCCGCA
AAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCT
TTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGAT
ACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACA
TTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGAC
ATTAACCTATAAAAAATAGGCGTATCACGAGGCCCTTTTCGTCTTCAAGAAT
TCCCGCCCGTAGCGCGCGCGCACCAGCCGGCATCGCACCCGAGCACCAGC
TCCCCCGTCGTCCAGATGCCACGGGCCACGTCGAGGGCCGACGGGGAGAA
ATACACGTACCTACCTGGGGATCTCAACAGGCCCGGGTGGCCAACCAGG
TCGTGGACGCGTTGTGCAGGTGCGTGATGTCCAGCTCCGTCTCGTGGGTGC
CGCCGGGCCCCAACCGGCGGTCCGGGGGGCGGTGTATCACGCGGCCCGCT
CGGGTGGCTCGCCGTCGCCACGTTGTCTCCCCGCGGGAACGTCAGGGCCT
CGGGGTACGGGACGGCCGAAAACGTTACCCAGGCCCGGGAACGCAGCAAC
ACGGAGGCGGCTGGATTGTGCAAGAGACCCTTAAGGGGGGCGACCGAGGG
GGGAGGCTGGGCGGTTCGGCTCGACCGTGGTGGGGGCGGGCAGGCTCGCGT
TCGGGGGCGGCGGAGCAGGTAGGTCTTCGGGATGTAAAGCAGCTGGCCG
GGGTCCCGCGGAAACTCGGCCGTGGTGACCAATAACAAAACAAAAGCGCTC
CTCGTACCAGCGAAGAAGGGGCAGAGATGCCGTAGTCAGGTTTAGTTCGT
CCGGCGGCGCCAGAAATCCGCGCGGTGGTTTTTTGGGGGTCCGGGGTGT
GGCAGCCACAGACGCCCGGTGTTCTGTCTCGCGCCAGTACATGCGGTCCAT
GCCAGGCCATCCAAAAACCATGGGTCTGTCTGCTCAGTCCAGTCGTGGA
CCTGACCCACAGCAACGCCAAAATAATAACCCACGAACCATAAACCA
TTCCCCATGGGGGACCCCGTCCCTAACCCACGGGGCCCGTGGCTATGGCA
GGGCTTGCCGCCCCGACGTTGGCTGCGAGCCCTGGGCCTTACCCGAAC
TGGGGGGTGGGGTGGGGAAAAGGAAGAAACGCGGGCGTATTGGCCCCAAT
GGGGTCTCGGTGGGGTATCGACAGAGTGCCAGCCCTGGGACCGAACCCCG
CGTTTATGAACAAACGACCCAACACCGTGCGTTTTTATTCTGTCTTTTTAT
TGCCGTCATAGCGCGGGTTCCTTCCGGTATTGTCTCCTTCCGTGTTTCAG
TTAGCCTCCCCCATCTCCCGGGGTGGGCGAAGAACTCCAGCATGAGATCC
CCGCGCTGGAGGATCATCCAGCCGGCGTCCCGGAAAACGATTCCGAAGCC
CAACCTTTCATAGAAGGCGGGCGGTGGAATCGAAATCTCGTGATGGCAGGT
TGGGCGTCGCTTGGTCCGTCAATTCGAACCCAGAGTCCCGCTCAGAAGA
ACTCGTCAAGAAGGCGATAGAAGNNN

TGGGGTCCTCTTTGGTCTGATGGAGAGAGGTTGGCACCAGGGTAAGTCGC
TGCTACATCACCCTGGTGTGTTTGTCTCAGCAGCTGGTGTAAATTTCTG
CCATCTGGGCTATTTCTGTAGAAAGCA AGAAGCTCTGCTGGTGGGCAGC
TCATCTCCCAGTGTGAAAAAGCAAAATGCAACGCATGCACCCTGCTATCC
ATGTGGBCCYAKCCCTCTCCATCAGCTGTTGAAGGAGAAATCTGCACTCA
GAAGAGATTGAATTGGGCTCAGATCTGGCTTGGGAAGATGATGATTCCAA
CCAGAGTCCAGGAGACTTTGGGGAATGCATGAATCCTATAGGAAAAATGGA
TAACCCCTTCATCCAAGAGCAAGCTGGCATGATGCTCTGGGGTGAAAACCC
ATAATGCCACCTGGTTTTTAAGGTTTGGGGTGGCTTACAATGTGCAGCTCT
GCTTCCGGCGAGGCACTGGGAGCCCTAAACCCATGGAGAGGTCAAACCAG
TGCTGGAGGTCAATTGTGGGCCAGCTGCAATGGGAGGTAGGCAATTATGG
ACATCGCTGAAGCCACCCACGCTCTGGGGAACCTTGGGTTTTACCTTTC
ACTGCACTTTAATGGGATTTCTCATCAATGTCTGCATGTTCTTGGCCACC
TGTTTAAAAATATAATAATAATAATTAAATCTTTTGCCCCACTGCGGGAT
GAGCAGCTGGTGGTTCCCAGCTCACAATAAACCACACTTGAGACTCCCTG
GAGAATTCGCTTTCTTTTGCAGCTGGTTCCATGTKGGGSYKTTAGCCC
CTCTGCAGCTCATAGGCTTTTCTTACAGCCTCTGCTCCACCTATTGCTG
AAAAGGGGGAAATTTGAGATGGATCCCATTTTGTGAACATCTCCCMACCT
GTGGGTAATGCTCAGACCTCTCAGCCCTGTGGGTTTAAATTTCTCTTTCTG
CAGCTTAATGGGTTGGGGATGTTCACTACTGCAATAATTAGTGATGGGAT
AGGGGAGGCAGGAGAGGATCCCGTCGACCGATGCCCTTGAGAGCCTTCAA
CCCAGTCAGCTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCGCGCAG
TTATGACTGTCTTCTTTATCATGCAACTCGTAGGACAGGTGCCGGCAGCG
CTCTGGGTCATTTTCCGGCGAGGACCGCTTTCGCTGGAGCGCGACGATGAT
CGGCCTGTGCTTGCGGTATTCGGAATCTTGCACGCCCTCGCTCAAGCCT
TCGTCACTGGTCCCGCCACCAACGTTTTCGGCGAGAAGCAGGCCATTATC
GCCGGCATGGCGGCCGACGCGCTGGGCTACGTCTTGCTGGCGTTTCGCGAC
GCGAGGCTGGATGGCCTTCCCCATTATGATCTTCTCGCTTCCGGCGGCAT
CGGGATGCCCGCGTTGCAGGCCATGCTGTCCAGGCAGGTAGATGACGACC
ATCAGGGACAGCTTCAAGGATCGCTCGCGGCTCTTACCAGCCTAACTTCG
ATCATTGGACCGCTGATCGTCACGGCGATTTATGCCGCTCGGCGAGCAC
ATGGAACGGGTTGGCATGGATTGTAGGCGCCGCCCTATACCTTGTCTGCC
TCCCCGCGTTGCGTCGCGGTGCATGGAGCCGGGCCACCTCGACCTGAATG
GAAGCCGGCGGCACCTCGCTAACGGATTCACTCAAGAAATTGGAGCC
AATCAATTCTTGCGGAGAACTGTGAATGCGCAAACCAACCCTTGGCAGAA
CATATCCATCGCGTCCGCCATCTCCAGCAGCCGCACGCGGCGCATCTCGG
GCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCA
CAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAA
GATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCG
ACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCCGGGAAGCGT
GGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGGTGTAGGTCTG
TTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTACGCCCCGACCGC
TGCGCCTTATCCGGTAACATCGTCTTGAGTCCAACCCGGTAAGACACGA
CTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGT
ATGTAGGCGGTGCTACAGAGTCTTGAAGTGGTGGCCTAACTACGGCTAC
ACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT
CGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTA
GCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGA
TCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAA
CGAAAACCTACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCT
TCACCTAGATCCTTTTAAATTAATAAATGAAGTTTTAAATCAATCTAAAGT

CCATAATATGCCATTAGAAGTAACACATCCATCAATGATATATCCATAGA
ATACAAGAGAACGGTCTACATTTACTTCAGATCCCATTTTTCAGGTTAACC
ATG.'AAAAAATACCCAAAGACTGAAT...ACCATTTCAGGGATCCCGTGTG
TAAAATCATGACTTCTGCTTTAATTATAAGAAAAATGAAATTCAGTGT
TTATTCTCTTTTAAGATGAACTCTCAACAGAAGTTGGTGAGTATTTTCT
GCCCTCCAGCAAAACCAAAGCATGCAGTTTGCAGTCTGTTTTGGATATAT
ATTGTACGTGGATATATAACCTGTATGTTATAACACCTCTGGTTTTCTTT
TCTCCTTCTTTTCTCAGAAAAACGAGAGAGAAGAATTGGTGAGTATCAA
ACTTCCCCCAGAAAGTGGACTTTGGTGTGTTGGGAAGATCCATACCACAA
CGTTGGTGCCAACTTAATGGAAATCCTTTGTTTTTCTTATGTTTTCA
GATGAACTCACTGCAGAGCTCGGTAAGTCGTGATTATAACTCATAACGAG
TTATAATGCTATTGTTATATATAATATACATATTATATATTGTTGCTATA
ATTCATAATAGAGCAAACAATCACAAGGCACAGAAATATGGGTTTGCTTT
GAGAGCCAAACCTTAGGAAGTGATAACACAATGGGAAGAGGACAATGACC
ATTTCTGTTGTTCTCTTTTTCAGAGCACTACAAGGCAAAAGCAAGTGAGT
GTCTCCTTCCTCATCTTCAGCACGTGAGAGATTTTGGGGGCTTTTGGGAC
GGCTATGGGGATTACACATAATAAAACAGAAGATGAGAAGACAGTTTGT
TAACTTGAATTCAAACCTGGTTTGAAATTGGTGAAATTACAGTATAAATAA
TCTCCCCAGTACCCAATTATACAATGGGATTAATTACAGCCTGCCCAGGA
AAGGAGCACTGAATTTTTTCTGCGTCCATCCAGCATGAAGTCCATCAGA
CTTAAGCTTACAGCTTAAAGAATGGTTCATTTTTTTTCATTTAACCCCTC
GTAAGTTAAAAGATGGACTTCAGCATCACAGAAGTAGCCCAAGAAATAGTC
AAAAAATGGGTCATGAATTTCCAGAGCACCCCCCACACTTTCCTTGGTG
AATAGGAAAACAAATATTAATACTAATTAAATTGGTTTTTTTTCTTTTA
GGAAGATGTTTTGAGGAACACAGTAAGTGCCCTTTTCTCCCTTCTTTAAG
CATCACTTTTCACTTTAAGTCTGCATCACAGTTAATAATCCATCTCCTTA
TTATGCATTTTTAGGGAGAGGCGAAGAAAAGTTGGGTAAAGTCATTTGGTT
AATTGGGTTTCTGCTTGCAGACCCCATCCAGGAGCTCATGTCCTCCTCT
AGTGTCTGCACTGTAGAAATATCCAGGTTAGACGTGTAGGTAGGAAATAC
TGGACCTGCGTGGAGGTATTGCAGACCCCATTTATGTGTAGGGGAAGCAG
AACATCAAACCTATTGAGCCTTGAGCTCCACGAAGACAAGCCACCTCTTA
GATTTCAAGCGAAGTCGAGCTGAATAGATTTAATTCTTTCTTTCCCATAG
TAAATGTGACTCTGGACCCAGAGACGGCCACCCTCGCCTCGTCTCTCC
AAGGACCAGAAGAGCGTCCGATGGGAATACAGCCTGCAGGAATCCCCCGA
CGGCCCCGAGCGCTTCGACGCCGATCCCTGCGTGCTGGGTTGTGAAACCT
TCACCTCTGGGAGGCACTGCTGGGTGGTGGATCTCACAGAAGGGCAGTAC
TGCGCCGTTGGGGTCAGCAGGGAGTCCCTGCCCAGGAAAGGAGCCGTCAG
CTTTAACCTGATGAAGGCATCTGGGCTGTGCAGCAATGGGGGTTCAAGA
ACAGAGCCCTCACCTCCCCTCCGACCCCACTGAACCTTCCACGGGTTCCT
AAAAAGATCCGCATCTCTCTGGACTACGAATGGGGCGAGGTGGCGTTTTT
TGATGTGGAGAACCAAATGCCCATCTTCACTTTTCTCTGACCTCCTTTG
GTGGGGAGCGGCTCCGGCCGTGGTTCTGGGTGGAGCTGGGCTCCCTCTCA
CTGCCCAGATAACCCCGGAATCCCTGGAGGTGCTGTGGAGGTGCCTTACA
GCAGCTCTTCCAGACCGGGGTGGAAAACTCTCAGGAAAAGCAGCATTA
AACCTCATTCTCCTCTTCCCAGTCAACCATTGTTCATGCAAAAGAAAGGA
AACCCATCCTCAATGTCATCAGCATCCTCCGTGTGTCATGTCTGGTGGCC
CCCATTGATGTATGGGGTGGCTCCTGTTGGTGTCTGGTGCCCCCTATTGA
CGTATGAGGTGGCCCCATTGACGTGAGGTGGCCCCCATTGACGTGAGGT
GGCCCCATTGACATATGGGGTGGCTCCTGTTGATGTCTGGTGCCCCCA
TTGACATGAGGTGGTCCCATGACCAGCCCTACCCTGGATCCAATGCCW
CMYGATTGCAGTTCCAACTCTAGGGACGTTAAACGACCCACAGAGAGGA

AATCACCTCAAAATGAGCCTGAATGTTTGCCTGAGGACTGAGCACAGCT
GGGCACTAATTCATCTTTATTTCTCTCTTATTTACAGAGGAACGCGATCT
GAAAATCAGTAAGTGCTGCCCCAAACCCATAGGGCTATGCTGGGCTTCAT
CCCCACAACATGAATTTTATAAATTAAATAAATAAATAAATAAATTT
TATATTTTATGTATTTGATATTAGCAGTATTTAAAAAAGAATAAAATA
ACTCAAGAATCTTAGGATCAATAGTAACACAATGATGCAACGTGGATACA
AAAGCAGTAATTCCTATTTCTTTGGGTTTTTATCCTTCAGGGGAACACG
AAGCAGAGATACGTGAGTGTTATTTTATATACTCTATAATGGAAAACCTT
TTTCTCTGTAATATAAAAAATAGGCTTTATTATTGAGGGGTTTTTTGGCT
TAACGCAAATGCGAAGTGCTTGAAATTCTACGTATGAAATAGAGGATTTT
CCATAGAGAAAAACAGCAATTTGGGGCTGGAATAAAAAGTTTCATTTCTT
GCTGAAAAGTGAATGAAAAGGGGGGGAAAAGAACATAAAAATTGAGTTTT
TTCCCTCATTAATCTGTCTATGAAATGGGTTGGGTTCTGAATGGTGATGT
CAACACCTCGTTTTGGGTTTCAGCCCAACATAATATGTGTCTGTCTTTAT
TTCTGTATCACTGGTGTTAAAGAGAGCTGTTTTGAACTAATATCTCTTTT
TTAATTACTTTTTCTTTTTCTTTTTCTTTCTTTTTCTCCCGTTTCTCT
CTGTTTTGCTTTAAGGGCGCCTCACTGAGCTGCTCGGTAAGTGCATTTCC
TTCCTTGCATCTGTSAAWMCAGCWATAACCVHAGGYCCTATTTTGGGGGG
GAAGGAGGGGATAAAACACAATAATGATGAAATCAGTGCTTTGGAAAGGG
TGCAATTATTATTTCTCCTGCAAATGAATACTTCCTTTTCCCTTTTGTTT
GCAGAGGACCGCGATTCCGATGTCCGTAAGTCTTTTGTGTGTCGGGAG
CTGTGAATCCTCCAATGGGAAATGCAGAAATTTAGAGTCTGCCCCAAAAA
TGACCTTTTTGAGGCTACAAGGGATGGGAAAATAAGGAGAAATGTCCTTA
TTTATTGATCTCCTTGTTTATGTGCAAACTGGGTGACTCTTCTCTGCCG
AACACGTTAGAAATAAGAACAACAAATGGGAGGAAATGGTATTTATTCAT
ATCTGTTGTTTTCTGTTTAATTTTTAGGAGAACAGGACATCCTCATTAG
TAAGTGGCACTTTGGATTGATAAGAAATGCAGCTCCTGGGGACGTTTGGG
TGCTGCGATTGCTGGCACTGCTGGGGCTTTGTGTTGTGGTGGAAGTGGAA
TTACTTCAAAAGAAGAGAAGAATGGAATTATCTGGAGAAAAAGGGGAATA
AATGGAACGTGTTTGGGAAAAGAAGGAGGAATAGAATGGAAATATTGGGGA
AAAAAGTGAAATAGAATGGAATTATTTCAAAAAAATGGAATGAAATTTA
GGGAGGGGGAAGGGGAAGTGGAAATGGAATTATTTGGGGGAGAAAAAGGGG
AAAATTGAATGACTGGGGGGGGAATGGGGAAATAGGATGGGAKTWTTTTA
AAAATACAGAATTGTGAAGGTTTCAGCCCATCTCAGAGAGTTTGGTATCC
TCGAGTTCCCCCTTTGCAACCCATTGAGCATCCTTGGGATGACACCAAT
TCTGTTTTCTCCTTTTCAAGGGAACTGTCAGAAGAGCTCGGTGAGTTAT
TTCCACTTCTTACATACAAACTGATTCTGGATAATCCTTTTGTGTGTTT
TCCTGCTTTGCCTCTTTGTGTTTTAAGAGGCAACTGCAGAAGGAATGGCA
CAAAGGGTGCAGAGGATCTTTGGGATAAATAACAGGGAAAACAGGGATGG
GATAGCAATGAGTTGGTGCAATAATCTATGGCACAAAAGGTGACGGCGTG
TTTCACATTTTGCTTTTTCTCTTCTTTTAGAGGAATTAAGGGGTGCGGA
AGTTGGTAAGTGAGATTCCTTTCCCTCTTCTCCCCAAAAGGATAAGGGGT
AATTTGGATTCTGATCTCTTTTTCTCCCTTTTTGTTCTAGAGGAGAGTG
TTCTGGAGAGGGGTGAGTATCATTCTCTTTCTACTGCTGCTTTTGACTGA
AGGAATCCCCCATAAGCATGCTGGTGGGATGGGAATTCTACATCTGATAC
ACAATTATTATCATTCTTCATTTTTTATACACAGAAATAGATAATTTTT
TTCCTTTCTTTTCTCTTTTCCCCCTTTTTTAGAGGAACATGATGCCAGAA
TTGGTACGTGTCCATCTCCCCCTGCTTTTGTGGTGTCTTCAAGAAGGCCA
ATGGGGTCATTTGGGATTGTTTGGGTTGAGGATTGGGTTCTTGATTGAAT
TTGGGGGAGGATTCAGGTGCCCAACACAACATCAGGTCCCATCTCATGT
TTTCCTATGGGCTTGATCCTTCTGTTGGATACCTAAGAATACCTGAAAT

TCCTGGGGGCAAGAGGGGCACCGCCGTGTAAAAAAWACATATATTTAACCA
TTTTTCCTTCTTTTTTTTCCCATTAGGGGACCGTGACTCAAAGCTCCGTG
AGTGCCACTCTCCTCCTGATTAAATCTGAGTGAAGATGTGGATTTTCT
CAGTGTGCTCCTACAATCTCACTTTTTCAGCACAGTTTTCCCCAAANTTT
GTGTTTCTCCACCCAACCCCTTACACTGATCCTAAATGGGTGTATTGCCT
GAATCAGTGGTTTTCTTCCCTATTTTNGATCTATCCNGTTTTATTCCAGT
ATATGTTTTATGACATAATTTTATGACATATTTTGTGTATGATGCCCA
TAGACCTTATTACCATTGCCTGCCCCTGTGTGGATCAGAAAATATATTTA
ATATAAACAGATATCTCTACTGACAGTGATTTCTGATGCACCCATGAAG
GAAAAGGATTTAAAATAAACTTTAATTTTTCCCTTTTTAGGCCAACTGAC
AGCAGAACTCGGTAAGSACMTKKYCKTCSYCATTCCCATAAAACAAATGA
AATTATGGATGGATGGATGGAAATTAACCAGTTAGTAGAGGTCAGCTTTG
CTCTAGGACGGTCTGAAAAGTGACCAAATCTGCTTTTACTCATTTTTCT
TCTTATTTTTTTGTAGCAAAGTGCGATGCAACGATCAGTAAGTGCTGCTG
CATGTGGGGGTACCTCCATCTTCGGGTCATTTTCTGCTGTTTCAGCATTG
AAAGGACATCAGAATTCTTAAATCCAACAAAATTGGGGTCACTCGAAAG
GAAATCTTTCAGATATGGGGGAAATCAGAGCCAAATTTTGAGGGGGGGA
GGGAAAATCTCAGGGGTGTTTCAGAAATCCAATGGGATCTGATGGTATTT
TCTGCTCTCAGGACTGTTTACAGTGGAACCTCGGTGAGTCCGTTTTCTTTT
TGTTTTTTTTTTCTAATTATTATTTATTAGTAGTATTATAAATCAATATT
ACTGTTGCTTATACATATTGTTGTACATTATATACATAATACATACATTA
TATACAGTATATAGTATACAGTAGTATATAATATTATGTATTATATATAT
ATAATGTATTATAATAATGTCATATCTAATATATGTCTGTATTAGATATA
ATGCATATATATTATTGTACTACAGTCATATTATAATACATTTACTTATA
TCTGCCTTTTTCCACACGTTTTCAATTGACCTGATTA AAAA ACTAAATCCTAA
AGGCAGAAGAAGATGAAAACCCCCAAATTAACACCAAATAATTGCAGCTA
TAGATCATATCTATCAAAAGCAAATTTGCCTTCAGTCCACATCACGAAAT
TAACAATAGAAAGGTTTAAATTTGGAACGTACAAACAATGACAAATAACC
CCAATGGCTTTTCTCTTCTTGCAGGAGAGCGTCACACCAAATAGGTAC
GTGAGGTGTTTGCTACCTTCGTTTGAAGGAAGAAATTGCATTAATAAAA
CCTCTGTCCAATATGAAGCCGGGGTCAAATTACTCATAAATCACCCTGA
TTGTCCATGAATTAACAGGGGAAAAAAAAGGCTAAACTTGAAAATAACATT
TTTTTCATCTCTCTTTTAAGGGGAACCTACTGCAGAAGTTGGTAAGTCTC
TTTCCCATCAGTTTAAGCAAAAATGGTTCATCAGATATATAATAATCCCT
TATTTCTGCTTGTTTTAGGGGACTACAACAGGAACTTCGTAAGTGCCT
TTAACTTCTCCCATTAAGAGTTAAACCTTTCAATATTTTTGATGCTTCAA
TGTGCTGAAGCCACCAAAAATGTGTTTYAATTGTAAAGGGGCTGAGCGTC
AAACCTGAACACTGCCATGTTGGGGGCTGAGATTTCGTGGGATTGGGTTT
TCAGTGTGAAAATGCCTCTGGGTTTCTGTGCCTGAGCTCAGGGAAACACG
ACCAGGGCTTCCCAGTAGGAATGAGACCCCAAATATTTCTACCTGGGGS
CTTTTCCCATTTGGGGAATTTATTCTGTWAATCCATATTTCTCCMCSTTTG
ARCGTYMCTCMTDMAATGTCACAATCTTGGCAATGTTGAGAAGATATATA
GATATCTATTTTAATACTGATTAATATGGAGGTGTTTGTGTTGGTCAGTG
ATGTCATCGGGAAAAGATCTGAGTCATTGAATCCCCATTTCTTTTCTCTT
TATTTTAAGGGAAACACGCAGCAGAACTTGGTAAGGGAATTCCCTCCCTG
GGTTTGTTCTCTTGTTTTCTCTTTTGGAGGGGGATTTTTTCTATGTCTT
CTTTCTATGTCTTCTTTCTATGTCTTCTTTCTATGTCTTCTTTCTATGC
TTCTTTCTATGTCTTCTTTCTATGTCTTCTTTCTATGTCTTCTTTCTATG
TCTTCTTTCTAAGTCTTCTTTCTATGTCTTCTTTCTATGTCTTCTTTCTY
AWGTCTTCTTTCKAWGGTCTTCTTTCTYAWGGTCTTCTYTTCTYWWCCTTCTT
TCTTTCTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTGGATTTTGAGCCAAAA

TAGAAGAATGGGATGCAAAAATCAGTGAGTGCCCTTTTTTCTCTCCCTT
CACGGTGAGGTATGGGTGTGGAGGACCTGAATTAATGTGAATTCCTCTGT
TTTAAGGGAAGCTAACAGAAGATTTTGTAGTCGCTTATTTTCTCGAT
CTGAGTGCATATTTCTACACCTTTACCATCAGTGATGACCAACGTGTGTA
TGCATTTCTCTTTATTCCATTTAGAAGAGAGCGACACAGAGCTCGGTGAG
TGCTTTGGGGTCTTATCAAGGTGGAAAGATGCCCCCTCTGTGCAACAGTGG
GGATTGGGAGAAGCCCTTCAGCTCTTCCATTTATCCACATCTGATACCCA
GATGGAGTCAGGATGCAGAACTGGAGGAGGAGGGCCAAAGCTTTGGGGCAT
TTTGGGGTTATTTTTGTTCTCTCGAGAGCTCCCAGGATTGACCCGTGTCCA
TTTCTGTGTTATTTCCAGAGGAATGTGACACAGAAGATGGTGAGTGTCTT
CCGTGAGAGGGCTCAGAGAAAAGACTTCCACCAAATCTCCCTCCTTTAATG
TATATTCTGATGTATTTATTTAAGGGGATCTCGCASCTGAGATCGGTAAG
TCGTGTGTGGTTATACACCCCTATKTGTGCCTCCCATCAAASAGGGCTCT
GTGCASCTTGAGTKGTGTTCCACAGGGTTTGTCTYCCCACTCTTCACACG
AATATGGGGGTAAAACCCAACAAAATGGCACAGAGGGATTGCAGAAAGGG
CGGGCGTTGGGTGGCGCTGTGTTCTGATCCAAGGGAGGGTGAAGCTCATG
AGAATGGTTCTTTCTTTCTTTTTTGAAGACAATCTGACTGCAGAGCTCG
GTGAGTGCTTCCCTTTCTCTCTGTTTCACTGTTGGGTTTTTAGG
GGGGAAAATGCTTATTCCCCCATAAACACACACATGTAACCCAACCTG
GGCTGGAAGAAGGGTCCAACGTTTATACTGCAGACTGCAATTATCATT
CCCAATTGGAAGGTGATTCCATCATGAACCATCCACCCATCACAGTGGA
TTCTGACAGTGTTTTCTCTGTGTTTTCCCTTTTCAAGAGAACGTGATAGGAA-
AATCAGTAAGTGCTTTTTTCTTCCAGAACTGATGGGAAGCGATGGGTTA
GGGTTAGGGTAAGGGTTAGGGTAAGGGTTAAGGTTAGGCTTGGGGAAAAA
TAAGTTAATACATTTTATTATGGCTTAGAATTGAACTAATGTTTCATCTA
TTTTCTTTGTTTTAAGGAAAGCTCACATCAGATCTTGGTAAGGGTTACTTC
CTTTAACTATCCTTAATTCTGCAACAGTGCTGGGTATAGAGTAGAAAAA
TATGCATGTGAAGGTGTATGTATGCACATGTTAATTCATTCTATTTATG
TACTCGTTAGTTGCTATATATGTATTAATTTATTCACATTATATATAT
TTGTATATATTTGCAAATATTTGTATGTATGTGTGTATGTGTGAAGAGAT
TGGGGTTTTCTCTGCTTGAAGAGGGGGGTGAATGACAGCAGGTGTCCTTA
ATAAGCCTTATTTTCAAACACTAACAGGGGAGAATTGGGATACACAGAA
ATAAAGCCTAAAAATGGGAAAAAGAAAAGAAATGAAATGGGTAAAAATATTG
AAAAGAACRAAAARTTTGGAGAAAAGAAATGACASTTTTGGTTGGGTGG
GGCTGCTCTGCATTTCTCCRCTTATTTTCTCCCTTTGCTTTTCAAGGTGATG
TTGACACAAAGCTCAGTGAGTGGAGCTGCTCTTCTGCCCCACATTTAAG
AGTATTTTTGGTATTTTTAAGACTGTTTAAGAATATTTGGACATTTCTG
TGGAAAATGGATTTCTGGTCTGTAAAAAAAACCTGGGGCTTATTTTTGAG
GACGGAATAAATGTCCCAAAAAAGGGGGATTTTGGCATCAATTGACTGGG
AGGTGAAAAATAAAGCAGTGATCTGAGCGTGTTGGGGCCAATGGATGAA
CCTCAATGATCATTGTGGTCCTTTTCAATCCAGGCCATTCTATGATTCTG
TGAAAGAAAAGAAGATAATTAACATTTAATTTTCTTCTTTCTCTTCAT
TCCAGAGGAACGCGACAGGAAAAATCAGTGAGTGTCACTTTTTTGGGGCCA
AAACCCTCTGATTTGGGGAAGGGATCCCTGATAGAAGTGGTTAATCCTGT
TGGTTTTTCCCTCCTTGCAGCCAAACTCTCAGCAGAAATACGTAAGTCCT
TTTCTCCCAATCTGAACTGTTTCTTTGTATTCTTAGACTTCCTTTTTT
TTTTTTTTCTGTTTTAATTAAAAATAATGCTTTTTTTTTGGTTGGTTTTTTT
TTTCCCTATTTGACAGGCAGACTGACTGCACTGCTGGGTGAGTGGTGCCA
TTAAATCCGTGTGTGGTTTTTGGGCTGAAAACCCCTTAAAAATGGGAACTCT
GCACCCAGACAGMYATSGTCWKWGCTTTKYWWCATTGTGAATAGAATAA
AAATGGGGGGAAATGGGCAAAATGAGCATTGCAAGKGAGCAGAGYTGCTG

CCCAAAATAGGCATGACTCAGCAAGCACCGTAGTGGGCATGATTTGCTTG
GGTGACCCCGTGGGTAAGGAGCCATTTGTTGGACACCACGATGTCGTTTT
TCACAGCCCTGTGAGCGCAGCGTCTTATTGCCCTCCAGACATTCCAAA
TTTGGAGAGCTCAAATGGCAAAGGTGAAATGGGCGTCAGCCCTCCGGGAT
GAAGGAATCTCTGCCGGGGTTTTCCGTTGGATCACAGCAGGAGGATTTGC
TTTCCTAAAGCATTAGAGTGACGTGGAGAGCCCAAATCGGACCCAGTGGC
CACATTCTCCCAAGGGAAAACCTTCGGGTGCCCTACGGTTCCTTTTCT
AGCATGATAACAAACTTCTTTTCCATCCGCCCATCCCCTTTTGGGTTTGG
AGGTTGACAAATCCCCACTGAAATTCCTATGTTGCACACATGTCCTTCAT
TCTTTAAGTAGGAGTTAGCAAAGGTTCCGCATTGACTTAATTCAGAGCGA
GATCAACAATTTTAGGCATTCTTTATGAACTTCACATTGTTTTATGCTGA
TCAGGAGCAAAAAACATACAGGAATAGGAGTGTGTCTGTAGGAGTGCTC
TGCATTTTCTTGCTCGTTTGGCTGATTAAGGAAGCTGGGAGGAAATGTTG
TGAAATAATCCCAAGTGATGAGAGACTGTGGGTATGGGAGGAGATGCCCT
CTGTCCTGGTGAGCAGTAGGGACAGAAGACCTGAGCTCATTTTCATATATC
TGTATATTAAGGCAATGCTAACCAGTGCTGTCGTGTTATTTGGGGCCAGG
AGTGGCTTCTGCCCCCGTTGGTGCCCATAAACCAGTGCTGCCCCATTKGG
GATTGGGGTTNTGTTGGCAGACAACATCCACCAACCAACCCATGGCTGAT
AGCAGAGAGGGCGACCAGGTCAACCCTCCATATATCTCTGCAGAAACCTGT
TCCTGTCTATACAGGGATCCCCATCCCTCCCCCAGCCCTCCTTCCATCCT
CGGCATTTGGGTTGGCTATAATTAGGCTCTGGGAACGTTCCCCTGCTGCC
AGCACAGCTGTCTGTCTGCAATGATCCTTCCAGCTCTCTGCGGACACGC
AAACCCTCCAGCAATCCTAAATACCCATTTCTCTGCACTCCTGGGACAAAC
TGGGAGCTGCCAAAAATCTCCAGCCCCCACAGACGTGACCATCACAGCA
CCAAGGAGCAGAGCAAGCGCAACGTGATTACGGTGCAGGTCCGGGGTAAGC
CTTTCTCTTTCTTCCCACAGCCCAGGATTTGGGGGATCCTATTGGCTCTA
TGGGATCTGGGAGATGCAGGAGAAATGTGATCCCTTTGCTGTAGCAAAAC
AACCTTTTAGAGTCCTGCACCTGAATCTGGCAGTACTGGAAAGCAGGAGA
GGGATTAAGAGTCCTTCTGCATTATCCTGCTCATAGGGAAATACAGCACA
GAAATCATTTGGGGCTGCTTCCTTTGCTTTCTTGGCACAAATTTAGGTCCT
CATTACAGCGTTTCTTTGACTGAGACCCCAATAGGATCTACAGGGGTAGA
ACAAAGCAGACAAAAAGTGATTGATGTTTCTATGCGATTTGTTGCCTTT
TCCCATTGAGATTTCTGCTTTTCTATGGGGCTTTTTGCTTTTTACAGC
TTTTTTTATTCACTGTAGTGAATAGAAATTTTTAGGGCTTTTAGGTCATT
GATGCTGTTATGAACACAGAGATGAACTCATAACACCTTCCTGGTGTGGT
TTGTCTATGGGATAGAAAGGAGCTCATGGTGTCTGTGGACAACTAACAGAG
GTGCCTGAGGGCTGGGCCCTCTTTGTGCCCCCTTCTGGGGGTCAGCAAAC
CCTTTTATTCAGATATAAATCCCCTCATCCACAATTTACCCAGTCTTCCC
AATGCAGACCCCAAAAAACATCCCAATGACAAAGTCCACGAACCTGAGAA
AAGCAGCAAAAAGCCTCCCAGCCCCAAATATTTATCCCTTATCCCATTTA
TTTCTATGGGCAAGCTATTCTAGGCATCAGGAAGGTGGGAGATTCCAGG
TCAGTTTGTTCCTAATTGTGATCTTTTAATGATGTTTCTCCCATCAGGTG
GACATTTGGAAGTGGTCTGACTGGGAAGAGGACGTGATGATGGCATCAG
GTAGAGCTCAGAAAGTGGTATTTATCAGCAAAGCAATTTTCCAGGTCTGT
TTTTTCCCATTTTCCCATATTTTTTTCTTATTACAGGGAAGAGGAACGCG
GATCTTGGTGAGTGATTTTCTTCCTTTTACCTTCAAAAAGTCCCTTTCCA
TGTGTAGAAATGGATATACGTACCCCCCACTGATACCCATTTCTTTGTT
CTGTCCTTATATTTATACTTCCCATATTTTGAACACATGAAAACAAAGC
CCACATTAAATAAATTCATAAACAGTGCAATTTTTGGACTATTATTTTCC
ATAGAAAAGTATTAAATCAGTGCAAGTGCCTCTGGAGGTGACTTCTGC
AGCACCCAAAGAGAGAGGGCTAGGGCTGAGTGCTCTCTGTCTCTCTTT

GAAAGAGGTTTATTTTTCACAGTGTGGAAACTCAGATCCGTTGCCTCACCT
GCACCGTGTATTTGCAGACACCCAAAGTGTTCCAGAGTTTGATGGTTTTG
TCCCTGGAGCCCGAAACGATCTGGCGGTTCTCGGAGGAGAAGGCGACGCT
CAGCACATCCTTGGTGTGGCCAACAAAGCGGCGGGTGGTGGTTCCTCTGC
AGGGACACCAGGAGGGTCGCACGGGAGGGACAAAGCTCAGCAAACCCCCA
TTAAATTAATTAACCTCCCTAAATTGAGGAGATCGTGCTGCAGTGCAT
AAATTCTTAATGAACACAACCTGATGGAAGCAGGAAGGAAGCTAAAACGGA
GTCATCTCCACATGGGTTGAGGAGTGGTGGTTCCTTCCTTCCTCCGAAC
AGGAACAAAAGGGTGCCAAAGCTTTTGATATAGGGTTGGAATAATCATGA
GGAGTTTAGGATATAAAACTCAGCTTCCGTGGACACACAGCAGCGTAAGT
GCTGAACGCTTTTGGAGGATTGGGGTAGTTCTGCTTCCTGAGGAGTTTCT
TCTCCTATAGTACTCCCAAAAATCACAGTGCAAGAAGAGCCGGTGCTGCT
CCAACCTCACCCCAAACCTCTGTACCCCAAAAATCACACCGAAGGAAAAGCC
TGCTTGCTCCAGTCTGTACCCACACAGCGATGGTGAAGGAAGAACCAAATC
CCCCCTGCTGCTCCACSTGCTTCTCTCCCATCATAATTGCAGGACGTGT
CCTCAGATCCCGGAGGATCAGCAGACTGTGTGAGGTGTAATCACTGGGAG
AGTGAGCTGAGGGAGGAACCGCTTGGTCTCCTCCCTCCAAGCATGATTTAC
CACCCAACCTGAGAGGAACCTACCTCATTTTCACGCTGTACCGCACACCT
CTCACCCACCCCAACACCCAAACAAAACACAGAGCCAGTTCTGCCCCAA
ACCCCAACCCCAAAGCCCTTTCAGTCCCCAGGACTCACGTGGTGAGGTCC
CACAGCCTCAAGGTGCCATCCCAGGAGCCCGACAGCGCAAACCTGCCCATC
GGAGGAGATGACCACATCGCTGACAAAGTGCGAGTGGCCGCGCAGGGCGC
GCTGCGGGATCCCGTAGTTGGTCTCATCTCGGGTCAGCTTCCACATGATG
ATGGTTTTGTCTGGGAAGGGGGAAAGGCAGCGGCCTCAGCTCCAACCCTT
CTCACATTCCCGTCTCTACTGGGCTTTATCTCCCTCATAGCAATGGGGGG
GTTACACAGAAGCACCGCACCCCTTCTCTCAGCCCCCAACCGCCTCCC
TACGTCTCATACACAGCAGCCTCCCCACCCTGCAGCTCTCTGTCCCCGA
GCCCTGCACCCCATTCATCACCTCCCTCCCCSAWGGTCCCCCCCCAGCCC
CCTCNTYTAYCAACKGACGGTGTCCCTTATTTCCACAGTCCCTCCATA
GGCCCCACAGTTCCCTGCCCCCCCCCACCCACAGTTCGSCCCCCCCCCG
CTCGGAMGAGGCCCGAACCCTCAAGGCGCGGCCCTCACCCCGCGACGSG
GAGAGAATCATGTCCGGGAACCTGCGGGGTGGTGGYGATCTGCGTCACCCA
CCCATTGTGGCCYTTMGGGTACCGSGGAGGGTCATCTGCTCCGTCATGG
CGGCGGCGGGGCGGAGGGATGGCGGCGGATTCAATAAAGGGCCCGGCCCG
GTCCGGTCTACCGCCCGYGATGGCCGCCAGCGCGGAAAGAGAAAGAGGG
AGGTGACTTCCGGCGGAAGCGGAAGTAGCCGCTGGGTTGTACGGCAAGAG
GGGCAACATGGCGGCGCGCATAGAGAGCACGCTGAATGGGGGAATGGGGC
TTTTGGAGGTGGGGAGGGAAGGTTGTTYTCTGCCGCTGCAGGRACACGAG
GTGCGGGCAGAGCACCTTCTTTAACATTTGKTATTATTTAACGTTTWACA
TTTAGCATTTTTATTATCCCTGTTGTGCCAGGACGGAGAAGAGCAGGGTG
TGCAGCCTGTGCTTATCACCTGCAGCTGTCCCTGCACCCACAGCCAACC
CAAGTTTGTGACGCCTGAGCAGGATCTGACCCAGGAAGGCAAACAGAAGG
TCTGAGTCTCTCCCTTTCTTTCCCATCCCTCCCACGCTGCAGTTTGG
GGGCTGTGACCCGTCCGCGTTGCTCAGTGCTCATTCCGATGAGCAGTGGC
TGATGGTGTGTTTACAAGTTTTTGGCATCCCTGTGGGTTCCACCCCCGT
TTTGTCTCACCAGCCTTTTTCTATCCGTCTTATCAGCAGATCATCCTTG
TTATTAGATCTGTCTTTTTCCAGTCACGGCTTTGCATTTTCACCTTGTT
TTACCACCTAACATCAAGCCTTTTGTCCCCATCTGATGATATTCATGCAG
ATAAATCCGTAAAGCAGGGAAGAATTAAATTCTGGCCCTTCTACACCCA
TTTAGGTTTAGATCTTTGCAGCATTACGCCAAGACGTGCTTCCAGAGCCA
GGAATAACGTGTCTTGATGTGCCAACACACCTTGAAATCCAGAAAATTGC

GGCCTGGCACTTTATTTAGGGCCACGTAGGCCGGGGAGGGTGCAAAAAAT
TGGGCAACTTCCACCTCTGAGGCTGCTCAGAGTGCAGCATCGCACCAGGC
CGCACCGGTGGGAAGCAGCCTTGTTCCCTTGCAGCTTAAGAGCTCTCT
GAGGTGGGGGTATTTATTTTCTCTTCCCTTTTCTCAGCTGCTGTTGAATT
TCCAGCTGAATCCTGTCCCACCAGAGAGACTCTGATTGCACCCTGTTGTG
TTTTACTTCTTTTTGTTGGTGGATTGGTATTTTTTTTTTCTGTTGGCGTT
ACAGAGCTAGTTCAAAATATTTTTGGCTAAAATAAGAATTAAATGGAGAT
CTAGTTTTTTGAAATGTCAAGAAATAATAATAATAATAATAAAGAATAAA
GAATAAAGTTTTAAAGCTGAGCCTCTCCCTTATTGAGAGCCCCCAGGGGA
CAGGAGTTGTGGTGCAGGCCCCCCAGTCTGCTGTAACTCCTGCTGGTAA
GATGTGACTTAAGCCTTGCATCGTTAATCTTAACTTAATTAGCAGTAATT
TGGATTGGGCTGCTTCCCTTCAGCAGCTTGTAAGGGGATAGAGGCTGCTG
GGTGAAGTGAAGCTCTGTGTTACCACCTCTCCTGCTCTCCCCACATGTTTT
TGGTGGTGGTGGTTGCTTCTTTTTGGCCACGGCTCTATCTCCCCAGGTGT
GCACTCACTGTGGGCTGCTACTGCTCCTGAAAGGGCTCAGGGAGACATTT
GAGTCCCTTCGTCCACACGTGGGAGGAGAGCACTGATGTCCCCATCCTTA
AAGTTGTGGGCACAGCCTTGGTGGCAAATCCAGAATGGGATATAATGCAG
CCATGAGCTCAACAGAGCGCTCTTTTATTGAGTTTTGTGCATAAAATCTG
TGTGTTGTTACCACATCCTCATCTGGTTCCAATGGTGAAGTTGCCACACCC
GGACGAGGTTATCTGTGTAGCCAGCAAACAGCGTCTGGGGAGAGAAATGG
AGGAAGTGGATCATGAAAAGATAGGAATCAGCCCTCGGTGTGAACGTAAA
AATCTCAGAAGGCAGCTCCCAAAGCGGAGGTGCTGGAGGAAGGTGGGAGT
TTTAAGGCTGCAGGAGGAGCAGTGAAAAGGGAAAGGAGAAGGGGATATTT
CTACCTGCCCATCTGCAGACCACGCCAGAGAGGTACACTGGGGAGGCTCA
GCTTTGCTGCTGGTGTGATCACCTCCTGCTTCAGCTCATCCACAATGAT
TTTGCCTTCCAGGTCTGTGCAGGACAGAAGAGAGCGTGAGGGACTAAGG
TCCTGCAGGGGAGACTGCTGTAGCCAAACCCAAACCATTCCTCAGAACCA
GGCTCAGGGTGCTCAGAAACAGCCTCTGGGTTTCCGCACAGGGATGCAGT
CAGATGGCATCGAAGTTTCATCACAGCAGAGTGGTGGCTGTGCCCCACAC
CACCTCCCAGTCCAGGGGATGACAGTGCCACCAGCATGACCCATCCCAC
GTAACCAAAAAGGGCTCTGCACCAAGGCATCTGTGGGGCAGGGCGAGGATT
TCGACCACAACCTCTGCCTCCCAAACCCAAACAGGATAAGGGAAAGTGATTCT
TTAGGAGGTAAATAGGGATGTCACATACCCAGATCTTGATGCTGGGGCCG
GTGGCAGCGCAGAGCCAGTAGCGGTTGGGGCTGAAGCACAGCGCATTGAT
GATGTCCCCTCCATCCAGCGTGTACAGGTGCTTGCCTTCATTACAGGTCCC
ACAGCATGGCCTGGCCGTCCTGGGGGGCAGCAAAGAGGAATCACAGCAAA
CCATCAAACCTGTGGCTTTGTTCCAGTTGTCCATCTAAAACCTTCCAGCT
TGGAACAGCACTTGATTTGTGACTGAGATGTGGGTGAGTTGCCACAGGA
CAGCAAGAGGCACATAACTGAGCTGTGAGAACAAACAGAATAAGCTGCAAT
TTGGCCTCAGCTTTCCCCCAGGGTGTACCTTGCCTCCAGAAGCACAGAGG
GAGCCATCAGGGGAGACAGTCACTGTGTTTACAGATATCCCGTGTGGCCGAT
GTGGTTTGTCTTCAGTTTGCAGTTAGCCAAGTTCCAAACCTAAATGAGGG
TAAACGTGACAGGCTCAGAAATATGGAGGAGAAAAAAAACAACCTCTCA
TGATCACTGCTCAAATATTCCCCAGAACGCCGCACAAACCCCAAAGGAGC
TGCTCCTCTCACCTTACCAGCTTGTCCCAGCCACAGGAGACAATGATGG
GGTTGCTGCTGTTGGGGGAGAAGCGCACACAGGAAACCCACTCAGAGTGG
CTCTCGTCCTGAGGAGAGGAACAGCATTGGGTTGAAAGCAATGAAAAGCA
TCCCCAGTCCGAGCTGCTGCATCCCACTGCTCCCTGAGCCCCCTCATAATT
GCAGGACGTGTCCTCAGACCCCCCCCCAGAAAGAAAGGTGAGCAGGCACTG
TGCACTTCTAATCATTAGGACGGAGCTGGGAGATGTGGATTACGGATCA
AAACCAACAAATCAAAGAGAAATGGGGGAATACGGACTCAGAAACAAGCA

GTGGTTTTTTGAGGGAAAATGCCGGGGGGGGGGGGGGGGGGGGGATATGCCC
TGAGAGATTTAGGGTCTGTTTTGGTAAGGAAAGCCTCCAGCAATGTGTGG
GCTGTGTCTTTGTTCTCTGTGGGGAAAGGAATCATCCAGGCTCAGTGTCTG
AGTTGTGGCTGATAAGAGGATTTATTGGGAGCAACGGTGGGATTGGTATC
AGTCATCCCTAATCCTTTTCCTTCTCTTTCCACCTTGCTGCCTCCTTCCC
ACAGGACATCAAGGGCACTTTTATCAGGTCAGTGACTTTGTTTGCATCTT
TTCACCTTTGAATAACTTTTCTTTTTTTTAAATGTCAAAAAGCATTGAGC
TTTTGTTTTAAATCCTGTGTGATGGGTACAGTTGGGGCCTGGTAATGCAG
GGGAAAGCTGTGTCTTAACCTTTTGGGTGATGGAAACTTCTGGCTGATGGG
GTGCAAATGGGATCTGGGGAACAACCTGGGAAAAGACTTGGGAACTTGGG
AAACAACCTCTGGGGCCATTTGGGAAAGGGGAAGGGTGGGGAGGAGATCTC
GGCCCTGATTTCTGGAAGCGTGGGTGTGCCCATGCAGACCTCATGCTATA
GCGAAACTCCTCACTCTGGAGAAACGATTCTCCCCATCCTGTCAGACAAA
TGGGCAGCGCTGGGAGTTCTCAGCCATGCTGGACGCACGTGGCTCTACCC
CAGCTCTGTCTGCTGGCTGAGGGAGGGTGGGGGAGGCTGGCTGCACCAGT
GCAACCAGTTTGGCCGATCCATGCGTTGCTCTGGTTTTTCCAGAGCTGCA
TGCAGGCCCGCCTCACTTCTTTTCTGCTGCTGAAATTCTCTGCTTTCCTCC
TTTCCCCCCCCACCCAAAAAAGATGTGAGAACATCAAATTCCAGGAGCCCCGA
GATGGTGTCTGGTGGACGTGGGGAAGAAATACCGCAACTATTTCTGCAAG
ATGTGGTGTGAGAAAGATGGAGAAAGCCTTCAGCAAAGTTCCACAGGGT
GAGAGAGTCCTCTTCTTCTACGTGGGATGGGGTTCCCTCCACTTGGGA
GGGATTTCTCCAGCTCTCTTGGGGTTCTCCTTCCATCTCTGTGCTCCCAT
GGTTTGCAGCCTGATGATCCTTTAGGAAAAGCAGCATCCCTCTGTTCTCT
CTGTGCTTTTCCCTTTTGCCTTGTCTGGGTTTTCCCTATTGTAGCTCC
TCCATAGAACTGGGGTTGATGTGGATCTGGATTCAATTATAAAGGAGGGAT
GACTGCCTCAAACCTCAGCATGGTGCAGATACGCAACCAGATGAGGATTTA
GGACTGGGGTGCAAGGGGGGAAAAAAGTGCCAGGTGACCCCCCTAACGACCC
CCGCTCTCTGCCCTTCCCTCCAGCTGACATCACGCTGGACCCGGACACCG
CTCACCTTCGCCTCAGCCTCTCCCTGGACCGCCGCAGCGTTAAGCTGGGA
GAACGACGCCAGGAGCTCCCCAACAACCCCAAACGCTTCGACTCCGATTA
CTGCGTCTTGGGCTCCCAGGGTTTACCACAGGCCGTCCTACTGCGGAGG
TAGAAGTCGGGGGCAAGAAAGGTTGGGCGGTGGGGGCTGCACGCGAGACG
GCTCGACGCAAGAAAAAACCATGGGGCCTCATCAAAAAGGGAGATCTG
GTGTGTTGGCACCAATGGGAAGAAGTACCAAGCGCTGACGGCCATGGAGC
AGATGGCTTTGTCACCCAGCGAGCGGCCCGGCGCTTCGGTGTCTACCTG
GACTATGAACGGGGTCAGCTTTGCTTCTACAACGCTGAGAGCATGACCCA
CATCCACACCTTCAACGCTTCCCTCCACGAGCGCATCTTCCCCTTTTTCC
GAATCCTGGCTAAGGGCACTCGTATCAAATCTGCACCTGATGGCCCTCC
AGCTTCTGATTTTTTTTTTCCCTTTTTTCCCCCTGCCTCATCCTTTGGGT
CCCACCTTGGGACCAGACGCTGCACTTGTTGTCTCGCACCTGCTTGCTCA
CAAGGCCTCTTCCCTCCTCTCTCCTGTCCCAGCCTCTGTCCACGTCCCAA
CTCTTCTCCGGGGTCGCGATCCCAGGCTGGTTTGGTTTGGAGAAGGGATC
CAATCTCCTTGCTGGAGGTTTTCCCTTACGCTCTTGGTGCTATGGGCTCC
CCTCTGCCTTTCCAGTCTCTCGCAGCAGCTTCCAGTGTGCTCTTCCCCG
TTTTGTTTTAAAGCCTGTGGTCGAGCTTTGCGTTGTTTGGCCTCTTTGGAT
GCAGAGCTCGAGCTGAGGATGCTGGGGTCTGTACATTGTGACACGAGCAC
TGCTTGTGCCCTCTTGGCCATTGCTTTCTGAAAGTCACTCAGATGCACCA
AGGAGCCTCATTTCTTTTTATTTTTCAGTTCTGGGGCACAACCCTCTGCC
CACCTCCCACCCAGCCACCATCTGGACCTCAAACCTTCCACGTTCTCCTA
TTCTGCCACTTGTCCACCTTCCCTTTTGTCTCTTCTCCCCCTCTGGGGG
TCTCAGCTCTCCCTCTGCCCCATCATTCCCTCGCCAACCATTTCTGTG

AGAGGAGGAAGATGAGCTGGGGGAGGAAGAGCTGGACGTGGAGCAGGAGG
AGGAGGAGGAGGATGGAGGCGGGGAGGAGGAGGAGGAGGACGACATGTGG
AGC JAGGAGGAAGAGGATGGAGAGC. G TGGGAAGGTACTGGGGGTTCGGTT
TGGGCCTGCCCTGTTGAGTGTCTTTATGGATGAGTGAGGGAATTGGGTGC
ACCCTCAGTCAGTTTGCAGATGATGCTAAGCTGGGGGGGTGTACTGATCT
GCCTGAGGGTAGGACGGCCCTACGGTGGGGTCTGGACTGGGCCCGATGGG
CTGAGGGCAATGGGGTGGAGTTCAGAAGGACCGAGTGCCTGGTTCTGCAC
TGAGGTCAACAACCCCATGCAGCTCTACCTGGGGTAGAGCGGCTGAAA
GCTGTGTGAGGGAAAAGGATTTGGGGGTGAATATGAGCCAGCAAGAGGCC
AAGAAGGCCCATGGCATCCTGGCTTGTATCAGAAATAGAGCAGCTAGTGG
GAGCAGGAAGTGACTGTCACCTCTGTACTGGCACACCTCAATGCTGCACCC
AGTTCTGGGTCCCCCTCTCACTACAAGAAAGACATTGAGGCCCAGTGAGGA
TGGTGGGGGTGGACTCAATGATCCCTGAGGTTTTTTTCCAACCTTGATGA
TTCTGTGATTCTCAGACCCCGTGGAAAGAGGAGCTGTGGGATGGAGTGGTG
CAGGGAGAACTCTACTTTGGGGACGATGATTATGATGAGGATGTGATGGA
GGAGGATGTGGAGGAAGAGGAGGAGGAGGAGGATGAAGCGCAGAGCCCTC
CGCCCCCTGTCTGCTGCCCGCCCTCGCCGCTGCAGACCTTCACCTGC
CCCCAGTGCCGCAAAACCTTTTTCCAGAGGAATTCAGACCCAACCTCCA
GTTGGCAAACATGGTGCAGATCATCCGGCAGCTCCACCCGACCCGCGAGC
GCCTCGCGCCGCCCCGCCGGCCCTCAGCCTCAGGGGGTCTTGGGGGGAAC
CCAGGGATCCTGGTGGCAACAGGAGGTGCGGGGGTGTCCGAATCTGTGCGA
GAAGCACCAGGAACCCCTGAAGCTGTTCTGTGAGGTGGATGAGCAGGCCGA
TCTGCGTGGTGTGCAGGGAGTCACGGAGCCACAAGCATCACAGTGTGTG
CCCCTGGAGGAAGTCGTGCAGGATTATAAGGTGGAGTTTGGGGAAGGGTC
ACGGTGGGATAGTGGGTGAGCTGGGGTTTGGGGAAGGGCTGTGGTGGAGA
AGGCGGGGTTTGAGGGAAGAGTTATGGGAGAGTGGAGGCTTGAAGGGAAA
GTGAGGTTGGGATCAAGCTAGGTTCTGCTTGTGAGCTGGTTGGGTGGA
GGCGTGGGAGGCTGGGAAACCACACACTGCAATGAGGAGGTGGAAGGGTC
TGGGTACCCATTTTCTGCTTAAAAACACCTTCCCAGCACAGTTCCTCAGA
GAAAGCAAAAGGGAAGTGGCGTGAAAGTTGGCTCTGAGGTTCCGTTTTCA
GCTCTGCCACCAAATTAGGGACA AAAAGAGGGCGATGACAGAGGGGATTGC
CCCAGGCAGGGTTTGCTGAGTTGTGTTTCCCTCCCTCAGTACAACTCCA
GAGCCATTTGGAGCCACTGAAGAAGAAGCTGGACGCGGTGCTGAAGCAGA
AGTCGAATGAGCAGGAGAAGATCACAGAGCTGAGGGTAAGAGCTGAAGGT
TTCTGTGCTTCATAGAATCATAACAGGAGAACCATCAGGGTTGGAAGAGAC
CACAAAGATCATCAGTTCCAACCATCACCGCTGCTGGGAGTGTGCCTTGG
TGGCTGAGCAAGGAGAGAGAAGCTTTGCTGCTGCTCTGAGCTCTCACGGA
GGCATCATATTCCTTTTCTGCAATTATTGGGCTGTGAGGGCTTGGAAC
GGTTTCCCAGTTGAATTAGAGCTTAATGAGAGCTTTGTGTGCCTCAGTGT
TGAGTGGGAATTGGTGGTTTGGGAGCTGGTATTCCTCATTTGAGTTGAGG
ATGCTCTACATCTCTAAACCTGTGCAGACTTTGCTCAGTTCTGTCTGTGG
TGCATTACAGGAGATGCGTAAGCTTATGGTGTGTGGTGAAACTGAGAGAAG
CATAGCACAGCAGCCCAAAAATGAGCTGATCTCTCACCTCCCCCTTCTGC
AGCAATTCCCCTAATGCTTTTCTCCTCTGTCAGGAAAAGATGAAGCTGG
AAATCAAGGAATTTGAGTCTGATTTTGAAGCTGCTCCACCAGTTCCTCATT
GGGGAGCACGTGCTGCTGCTGCACCAGCTGGAGGAGCGCTACGAGAGCCT
GCTGGCCCGGCAGAGCAGCAACATCAGCCAGCTGGAGGAGCAGAGTGCAG
CCCTTAGCCGCTTATCACGGAGGCAGAAGATAAGAGCAAGCAGGACGGG
CTACAGCTGCTCAAGGTCTTCTTCCATCCCTTTCTTGTCTTTATGGCAA
AGCGATAGCACGATGGTGGGAATAATGCTCCAGAAAGCTTCTGTGTCA TG
AGAGAGTGCCTTTAGTTGGTGGGCTGGGTGCTTCTCCACCCCTCCTTGTG

AAGATAAGCACTTGGTGAGATTTCCCTCATAAACACCCCCAAAACGGCGGC
CCTGGGGTGTGTTTCTGTATTAAAGAGCCCTCAGTGGAATGGTTTTTGA
GGG JTGTTGGTCGAAGAGCAAAGCATCAAAGGAAGGAGAGGGCAGTAATGT
TGCAAAGGGCTGACGGCGGTGGTTGCAAAGAGGGAGGATGGGGGGGGATG
CGCCAAGCAAAGGGTGGCGTGGGTTACCCGCAGGGATGCACTGCGCCCT
TGGCTCCGGGTTTTGGGACCGTACCTTGTACTCCTGGGCCCGCTGGTGGG
CAGGGAGCACAGCGTGGGAGCGGTGCGCCTGGGACGCGTCGCACTGCGCG
CAGATAGGCTCTTGGTCCTCTGTGCAGAAGAGCTTCAGAGCCTCGCGGTG
CTGCTTGACCAACCCGAGGAATGCAAACCTCAGCTGCCGGGCGATGCTGG
CGATATTTGCCAGCTCTCTGCTGGGGCGGAAATTTTTGTGCAACGCCGTT
TTCCTGCACTGCGGACAGGGGAAATTTCCCTCCAGCCCTTCCCAGCAGCG
GGCGATGCACTCCCGGCAGAAATTGTGGCCGCAGGGGATGGAGACGGGAT
CCTGGAAGTAACCCAGGCAGATGGAGCAGGAGGCTTCGCTCTGCAGGCTG
TCCAAGGGGCTCTGCGTGGCCATGGGCTTCCTGCTGGGCTCCGATCCGCA
GAGGGAATAGGGGACCTTTCCTCCTTATCTCCTCGCTGATAGGAGAAATC
CGGCCCCGGAGGCTGAGCCTGAGCCAAACAGGGCTGGGAGAGCTCAGCCC
ATAGGGGATGCTGGTGGGAATGGGGGCAGCTCGCGGCTCCCCAGCACGGA
GTCACCAAACCTGGGGGGATCTGGGGGAAATTCGGAGGAAAAATCAGATTT
TGTCTCTCCTCGAGCAGCAAAGAGGGCAGGGGAGGCGATTTTCCTTC
TGTGCGATCACTGTAAGGAATTTCCAAAGAAAACGCATGGAGGTCTGCTT
GTTGGGATGGAATATAGACGTATATTGGAATAAATACAGGAAGACGTTGG
AACATGGGAAGGCACTGAGATATAAGCGTGCTGTGTTGGATATGACTCTG
CTCGACTAAAGTGAAGGTGGTTTTAATAGCACTGCTCAGAGCCAGGCGGG
TTTTGGTGTGTTTGGGGGGAATTACGTGGGTTTGGAATTGGGAAATATG
AGACGGAAAAATAAGAATAATGGAAGCGCCCAACGTGGGGCTCGAACCCA
CGACCCTGAGATTAAGAGTCTCATGCTCTACCGACTGAGCTAGCCGGGCT
GATGGGCACGCACCCTTCTAAGCAATACTTCATGGTGATCCTGCGGAGGG
GTGCTAATAATTCTACCTAATTATTTTGTTAATTATCCCGGTAATTATGG
GTTCTGAGCAATCGCGAATCCACGGGGAAGAGCTGCATGGGGAAAAAGCA
CCTATCCCTACGGGAATAGCCGGGAAGTGCCTCGGAGGTTTGCAGATAGG
GGCCGCTCCGGGCCGCTCCCGCGCCGTTCCGGTGAGCACAGAGTGCAGC
GGGTGACAAAATGAAGGGAAAAATGTAAACTGATGCTCCCGAATCGAGG
CTCGAACCGCCAATTGTCCGACTGACAGCCGCGCGCTCTACCGACTGAGCT
ACCCGGAGACAGAGCAGCCGGAAGTCACGCCCCCGTAGAGCGCCACCCC
GTTGCCTAGTGACAGGAGCGCCGCTTCCGGTCAAGTGATGAGCGGAGGGG
GCGTGGCTTGTGTCAGATAGGACGGAAGTTCCGGTCAGGTGGTACTGGAA
AGGGGGCGTGGCTTGCGGCAAAGGGGACGGAAAGCGGAAGTGCTGCCGTT
GGTTGGCGGAGTTCGCACCATAGAAGAACGACGGCGGCGGTGGGAGGGCG
GGAGGTAGAGCGGTCCCCGGGGAGAGTGCTGACGGGAGCGGCGAGGCCCG
AGGAGGGAGCGGAGCTTACGGGGGAGTGCGGAGCCTCGAGGCGGGTCCCA
GCGCTTCGCTGTGGGGCAGGAGAAAGGCTTCGGGGCAGGAGGAAGAGGGC
CTCCSGGCCWSSCSATGGAGGCGGTGGGCGACGATGGGGCGTCGTCGGGG
CGGCTGAACCCGGTGGAGACGCTGCAGGAGGAGGCGATCTGCGCCATCTG
CCTGGACTIONTTCGTGGAGCCGGTGTGATCGGCTGCGGGCACAACCTCT
GCCGGGTGTGCATCGCGCAGCTGTGGGGTGGAGGAGAGGCTGAGGTGGAG
GAGAGCGGCGGGGCCGCGCGTGGAGGAGGAAGAGGAAGAGCTGGAGGA

AACACGCAGGGCTCGAAGCTGAACCTCTCGGGGTTCTCGGGGAGGTCCTG
TGGCACCAGTTGGCCCCGGGCTTATTTTCGGTCTTCAGAGAGATGGAGGT
TGGGGTGAGCGGTGGTGGGGTCCATGGTGACGTTGGCTGTGGGACATGAG
GGGGAATGGAGGTAGGATTTAGGCTTGGGGGGAGCTGGAGAGGTTCTCT
TCCTTCTGTCTTTTCTCTGGGTGCTTTTGGACATGGGCTGGTGGTGGT
GTGGGTTGATGGTTGGGCTGGGTGATCTTTGGGGTCTTTTCCAACCTTTG
TGATTCTATGGGGTGTGTGGGGCTCCACCAGCCTCAGTGTCCCCCAGTAG
AGATGTAGGAGAATGGGGAGAGGACAAATTTTAGGGCAGCATAATGCGGG
AGGGACAAAGACATGGGAAGGGGACAGCTTGACATTACGGAGGGGAAGG
GGAAGCACAAACACTGTTAGGTTTTGCCTTGAATCTGTTACTGGCTTTGT
AGGACCACCAGCATCAGGATGCTGTCCCCATTCCCTCCCTTCCCTGTGGG
ACTGCGTTGTTTTTTTCCAAGAAAACCACTCCCCACCCACATCCACCAC
TGCTGACATACCTGGCTCTTGCAATTGAAACATCAGGCTGTCTGAAAAGG
AGAACAAATTCATGCAATTGGGTTTATGCTTCAGGAAAAGGGGCTGGGAG
ATGGGGAAGGGAAACCATGGGGGTCTGGGGGCTTCGCAGTGCAAAAGCTC
TGGGTTTACTGCAAGAGCCCCACGACCCTCCCAGACCTGGAGGAGACCCC
GACCCCATTCAGTACCTTGGCACTTCTGCAGCGTCAGTCTCACCAGGACG
TTCTTCTGAAGGAAGTCCTCCAACCTTCTTCCAGAGTGGGGGAAATCTC
TGCTGGAGGGCTGAACTTCATCATCTCACAGCTGCAAAGAGAGGAGAAGG
GTGGGGATGGGGGGACTGTTGCGTTGGTTGGTTGGCTGTTCATTTTATTC
TCAATAGGAGAAGCTATGGGGTGAGGATATTTGCACAGGGACGAAATCCC
TTTCCCCCTGGGATCCCTCTGCCTTGACGCCCTCCCCCAGGGTGCCATC
CAAAAATCAGGGTGACAATAGGAAGGAGCCATGTTACCTATTCAAGAGCC
TCCTGATGTCCTAAAGGTGGGAGGAGAGAGGAGAGATGGATCAGAAGAGG
AGCACCAAGGGCTGCCCCCTTCGTATGGCAATGCACAGCAAAGACCACCT
GCCCACGGTGTGATCCCCCCCAGCAGCAACACAGGGAGCTCCCATGGGGT
TGAGTTTGGGTTCTCAGGGTTTGCTCTGTCCCCCATTTCCCACCACCCC
TTTGGGTTCTCACCAGCAGGAATTTGCTGTGGGGCTGCTGGAATTTGCC
TCCATCTCCCAGATCAGGGTGTCAAGGTGGGACATCTCCTCCATCACCTT
CGTCACCGCATCCTCCTGTACTTTGGTGACGGCTCTGTCCAGGTCTGCCA
GCTGGACCAGCAGGAAGCGCTCCTTCTCCTTCAGAAATCGCTGCAACTGC
TCGAATTCACACACTATCCTCTTCCCTTCTTCTTGGTTTTCTCCTGTTG
GGATGAGGGAGAAAGCCAATGGGGTGAATAGAGGCAGGAAGACCCCCC
TGGGGTCTCAGGATGCCGTGTTCTGGGGGATATCCAACCAAAACCAATGG
GGATGTAACACCAATGCCAATGGGAGCACAACACTAATGCCAATGGGAAT
TTATCACCAGTGCCAATGGGAACGTAACAACAGCGCCAATGGGAACGTAA
CACCAGTGCCAGTGGGAATTTATCACCAGTGCCAATGGGAACTTAACATC
AAAAAGCCAAAGATCATCTTGCTGGGCATTTGGGAGCAGCAGGAATTTT
CAGGAGTTTTATCCCAAAAGCAAAACCAAAAGGAGGGGGTAGGAGATGAGC
TCTGTATGAGGGATATTTACAGAGTTTAGGAGGATCTGCTACGTTATCTC
TTTAACACAGGGGTTCTGCGTAACCCAGCTGATAAACACAGCCTTAGC
GCTTTCCCAGCCCAGCTGCGAGCCAAAAATGCATGATCTGCCCCAAAAT
ACACCAAAACAAACAGGACAGGGCGGAGGGGAAGGCAGACACCTCCCCTG
CTGCACCCACCAATAACAAGCCCGTCCTTCCACCAGTCCTTCTGCTTTCC
AGGTACTTTTTCCCTCTCCTCCTTTGAAGCCTGGAGGCGAGCCTGAATTC
TTCCTGTGCCAAAAGAAGAAAGGCGGAAAGCCTGTTTTCCCACTTAAACT
GCTTCTGTGATGAGGAGGGCTTTGCTAAAGCCTGGAATCCTCTGCAAG
GTGCAGAGCTGGGCAGAGGGAAGCTCTGTGAGCACGGTGTGCTGCTCTGG
AGCTCTGTGCAAGCTGGGAGTATTTTGCAGAGAGAAAAGAGGGGAGAAGG
GAAGGAAAAACACGAACCTTGCTGCAACGTAAGAGAAAACGCTGCAAAAG
TGCAACAAAAAATCAGCACTGACAGCTGCGCAAGGAGGTGTGGAAGGGC

CAAGAAGGGCTCTGTGTTTTCTGCCTGGAATCTGAGCCCTCCCTACTGG
GGCTCAGCTTTTCTTCTGATGCAGAAAGTGGAAAATAAAGAGCAGTGGGA
CTGGAAATACCAGGGGGGACTCATGAC TGGCATCCCCACTGGAGGAGCT
CAATGGTGAGCTGGAATCCTTGCTAAGTTTTATCGAATGTGGGGGACAGG
AGGAAGAAATCAAACCTCAAAAAGTCATGAACAGGTGGCTGTGAATTCGGG
GCAGAAAGCTGAGGGGCCCTAAAAGCACAGGAGGCAAAAAGGATGGAGAGA
AACGACCCTACTGATGACACATCGCTGCCAGCAGCTGACACCTACCAGA
TCCTCCAGGTTTGGGCACTCCAGGGCGCTCTTCTTCCTCGGAGACTTTCT
CTCTCCTCCTTTGGAAACCCCTGATATCCCTCTGAGTTTCTTCCCCAGTG
AACCACAGAACCTGTTGTTTTAGCCCTTTGATGGGGTTGGGGTTTTCC
CTTCTGTTCTTCCCAGTCTGGGGTAGAGCTATGGGATGGCTGCGTTGA
GCCTGCAGGTCTGCTCCTGGTGGCACCCCTTGGCAGGGCGTGCTGGGAGCT
CTGGGTTTGTCTTTGTCTTTCTCCAGTTCTTGTCCCGGGGAGATGCT
GAACAATGTCACTTTGCAGATTTTGTGAGCTTCTTTTAGGATCGAGCCA
TCGGGAGTGGGGTTAGGGGGTGTATATGGGGAAACCATAAGGAAATAGGG
AAGGAGATGCACAGCCGGATCCTTGTGGGGATGTGGAGGAGCACAAGTGA
GGATCTTTGGGATTTGAGTGCTCTCTCAGCCCAGCACTAACACAGAGCAC
TCACAGCCCTGGCTCTGAGCTCTCGAGGAAACATTTCCAACCATTTCTGC
CCCACTGTCTTGTGTTGAGCCCCATGGCCAAATACACATGCCTAGAAAA
TAAAGCCATGCATTACATATGTATTTAATTTTTGCGTGGCAACCACTGAG
ACCCAAGTGGAGGAGATAACTGCCATTCACTTGGGCAGGTTTGCAGGGGT
GAACTGCACTTCCAGCAAACCCCTCCCTGTTGGGAAGAGCCACAGGGATGG
ATGGCACTCTGGGAGCTGAAGAACTGGAAGCAAACCTCCCTGCAACCGCTC
CCCTGGGGCACAGAGCCTTTTCATCCCAAATAAAGGCGTCCATCATTGAGC
AAATGAGTCACACCGTTGGGCAAACGACTTGCATTGCATCCCGAAAAGCA
TTAATTGCAGAGCCTGGAAAAGTGTGCTGGGCTGGAAACATCTGCATTGCA
GATCTACGGAGCAGAATAGACCCTGAACAGATCCTTCACCCAAATTCCCC
AGCAGGTGGGACCAAATGGCAGCGATGCGTGGGGCTGAGGAAAGATACCA
ACACATCAAAGAGCAATATTGAAATTTAGCTGTAGGTTTGACCTTTGGA
GGTGGTGAGGTGGGGCTTTGTGATGGGATACCCACTCATATCGCATCTGC
TATTCTGAGCCTGATGTGCGCTGCTCCCTCCCACCCTCTTTTAGTTCTC
TTCTTGGTTCTACAATCACCACCTGTGTGTATTTTGGTGCTGCCTGTT
CTCTTTTGGGCTTTCTCAGAAGAAAATGGGTTTTTGGGGAATCCATTCA
GGTGAGTCCTCACCCCAAGCAGCTCTTCTTCACTTTGTTGGCCCAAAGCT
GACCCAGAGCCATACACCCAAAGCAAACCCAGAGCCGTACACCCATAATG
AGGCAGGAAGTGGAGTGTGCAGAGCACATCTTTTAATTAATAAATACTAT
CAGAAACGTAGGCAGAGACCAGCTCCCCACACCAGGCGTTGCTATTTGCA
GTGAAAGGCCGCATACCTTTGCAGGACACCCCAAGATCTGCCCCACGATTG
ATGTCAAATAGATGCATAAATTTCTTCCAAGTCTTCAGTGCTCTCTGGT
GGTTTCCCCACCCTGCAGAGGGACCGCCCCGGGGCTCCCAATGGGGACAG
ACACAGGGCAGAGCAGCGGGTCCCCTTGGCACATTGCTCCAAGCAACCAC
AGCACACATCCCATCAGATGCCCTTTTATAAAGGACATCTCAAGGACAG
ATCTTTAGGGGAGATCTAAACCCAAACCAATCCAAATGGGACATCAGCTG
CCCACTCGTGGACTGCTCCTCTGAGGGGGGATTTTGGGTGATCTCTTGCA
AGCGAGCCCCCAGCCCTATCTTGAACAAGGGGAGGACCTTCTCCCCATTG
AACAAAGCCCTGGTGTACACCAAGATGGGGGTGTCATCATCCGAGCTGAA
GAATGCCACCCGACCCCTTCGTAGTCCAGGGAGACCCGAATCCTCCTGG
GAAGTGCATTACAGACGTAGGTTGGCACGGGGAGACGTGAGGGAGTGGTAG
GCCTCCAGCGCCAGACACCCTCTTTGGGGCTGAAGCTCATGGGTCCCTT
CCTCTTCATCGAAGCCCGGGCCACCCCAAGGGCCACACCCCCCCTGTC
CCACCTCCACCTCCAGAAATGCCTCCCCGAGGTGAAGCCCTGGCAGCCC

ACAAAGATGAGCCCCCTTCATCAGCATCAAAAAATGCCACCGTCCCTCC
AGCGTAGTCCAAGTGGACGCTGACCCCTCTGGGCACCCAGCGCAGAGCTA
ACAGGGTCACCTTGTGGGTGGTGAAGTCCCCGGACCTGTCCCCCCCATTTC
TCCACCCCCCAAATCCCCCCTTTGGGACAGAGGCTGAGTTGACCCTTCCG
AGGGATGGATTCTCGGGCCACACCGATGGCCCAGTCCCCTTCATCCCCCA
CTTCCACCTCCCAGCAGTGCCGGCCGGCAGAGAAGCTTTGGTGGCCCCAAA
ACAAAGGGCCAGTAGGCGAATCTTTTCGGGGTTATCAGGAAGGTCCTGTTG
TCCTTCCCCACGTTTTCACACTCTTTTCGGTCTTCGGAGAGGATGAGGTCAG
GGTGAGCGGTGTCGGGGTCCAGGGTGATGCTGGCTGTGGGGTGGAGAGGA
TGAGGAGTGTAAGGTTTGGGTCTCGGTGCTGAGGCCATGAGGATGCGGA
GAGCTTGGATCTCCAGCACTAAAGGAGTTGGATGTGCTCTAGATGGCCCC
ACCTGAGTAGGGTTGTAGGGTGGGACCGTCCCTTCCAACCTCAGCCATTC
TGTGGGGCCATGGGTGGCATCGGAAGGGTAAAAAGTACCAAAGAAGAAA
GTAAAAAGGTGAGAGGTGGAAACCCCTCTCATGTGCCCGTGCTATATGAC
AATAAAAGTGTTTTGAGCCCCCAGAATGCCCAGAAATAAAGGCGTTTCTG
CAGACCTTCTGTTCCATTGGTCAAAAGAAATGGTGAGGGGAATAAAAAATG
GAAGGAAGGAGATCTATGGGATATTACCTGCAAAGTCTGCAGTGCTTCAT
CTCCTAGACCAACCCGGACCAAGTTCAGCCAACCCCATGGTTTAAAAACA
GAGCTGAAATCTGAAGGCAGGGATAATGAATGAGTTCAACCCGCTCACCA
TATTTGTTTATGGGAAATGGATATTTATCAAGGCGAGGGATCTGCCCTGG
GGCCATCATCCCAAATTACAGCCAGACTCGGCCTGCAGGGTGAAGAAAAC-
TTGTTTGGCTGCCCTGATTTTTGTGTATTCTCCCTCGGCATCTATTTTT
GTCCATTTGGGTACAGCCTATGGGTCCAGGCGCGCCTCCATCTAACAGGT
AATGCGGCTTTAGGTTCTCATGCTCAGCAAAAGGCACTTTTAGGAAAGGT
GAAGCTGGAGGGGTGCAGAGCCGGAGAGCAGCCCGTCTTACCCCTGAG
CACTTCTCAGGAATTACAGCAAAACGTGTAATTAAGAGTGGCAAACGGGG
TATCGAGTCCTTCGGGTCTCAATTATTTTCTGAGTGGGAATAACCCGTT
GCTCTTCCATCTCTCTGCATTATTCTGCTGCAGAACGAGTGATGGGCTGC
TGGTTTTTACCAAAAATACCACCATTTCCACCCGAAACCCTTCTGAGTAC
CTTGAAGCCTCTTCAGGGTTTCTTTCAGAGCACCGTTCCTCCATGAGGAA
TGGCACAGCCTCTCCTCCGGCCCTGGAGAAGCGCCCGCTGGCAGCTGGAA
GGTCACTTTTCCACACCTGGAGGGGAAATAAATGCATTTTCAGGTGGTTG
TATCACAGAGCATGCCATCACTTCAGGACAGCAGAGGCCAGCACACGGCG
GCCATCCCCAAAATACCCTTCAGGGCTCGCAGTTCCCTTGAGCAGAAGA
GCATTCAATTGATGAGCTTTCTCCTCCATGGTCACTGCCTGATGCAAAGCT
CACAGAACAGCTTTTCAGAGAGGCCACATACCTGGTGATGGGGCTTTTCA
CATCCTGGGGACAGAAGAGAGGAGGGGGAGAGGAAACTCAGGTCAGTGCA
TGACCCATTTTGTCTTTAAAGTATGGAAAATTGAGCTGTTTGAGTGGGGG
TGGACCTCTTGGGTCTTCCAACATGTGCCCAATTTTGACTTTAAGTCATA
GAAAAAGTGAATTGTTTGACTGGGGATGGATCTGTTGGGTCTTTCAACAC
ATGGTCCATTTTGTCTTTAAATCATAGAAATAAAGAATTGTTTGACCAGA
GATGGACCTCTGGGGTCTTCTCCACGAGGAAGGTGAACCAACTGAGGAG
CATCCATGCACGGCAATGAATCCTGCAGATCCACCCCACTGCTGCTCTCC
CAACCCAGCCGTGGATTTCCCTCTTAAACAGACCCCATGAGGACCTTC
TGCAGTAAGGTGAAAATACTGGGAATACTGAGATGAGGATAAAACGGTGG
GGGGAAAGAGGAGGCTGCAAACCTCCATCTCCTCATTGTGGTGGGGGTTT
CAGGCTGATGGAACGGCATAAAATGGGAGGAAACACCCAATTAAGGCAC
CATGCAATTGGTCGGGGTGGGGAGGACATCCCTAAAGGACTTTTCCCTT
GAAAAAGCTTCCCTGGAGGAATTAATCACCAGCTGCTGGCTCTTCTCTC
CCTGTGCTTTTCGTATCCAGCGGGGAAATCTCCTCCGAGTGCTTGGCGGTG
CTTTTCTGCCTCTTCTCAATCTCATTTTTTCAGGTCTTCCAGCTGCCAGAG

TCGATCAACGTTTCAATGTTGGTATCAACACCAGGTTTAACTTTGAACTT
ATCGGCACTGACGGTTACCTTGTCTGCGCTGGCTCATCACGCTGGATAC
CAACGCTGATGTTGTAGATATTGGTCAJCGGCTGAGGTGTTTCGATTGCC
GCTGCGTGGATAGCACCATTGCGATAGCGGCGTCCTTGATGAATGACAC
TCCATTGCGAATAAGTTCGAAGGAGACGGTGTACGAATGCGCTGGTCCA
GCTCGTCGATTGCCTTTTGTGCAGCAGAGGTATCAATCTCAACGCCAAGC
GTCATCGAAGCGCAATATTGCTGCTCACCAAAACGCGTATTGACCAGGTG
TTCAACGGCAAATTTCTGCCCTTCTGATGTCAGAAAGGTAAAGTGATTTT
CTTTCTGGTATTCAGTTGCTGTGTGTCTGGTTTCAGCAAAACCAAGCTCG
CGCAATTTCGGCTGTGCCAGATTTAGAAGGCAGATCACCAGACAGCAACGC
GCCACGGAAAAACAGCGCATAACAGAACATCCGTCGCCGCGCCGGACAACG
TGATAATTTTATGACCCATGATTTATTTCTTTTAGACGTGAGCCTGTGC
CACAGCAAAGCCGCCGAAAGTTAACGGTTTGGCCAGGCTCACAACCTGAAA
GACTTTCTACGGTGTGCGCGTGCATGCGCGTAGAAGACTGATTTATCAA
CCTGTCTTTATATCAGGATTCATTACCTGACTATTTGTGGGTAAAGTTTCG
TAGTGCGCTGATCGTGCAAAATGATTTTAGTTGGGAACAGTTCGCAACTC
TGTCCTATAAAAAATCAGCATATTCCCATCTATCCCATATCCAGCGCATTG
ACCATCGGGATACTGAAGGGAGATTCCATCATCTCTTAGAAAGATCACCA
TCTCTTTTGTTCATTTGCATATAGCTACCTGGAGGATTTATGAATACA
AGGATTTTCATGGACTATTACCATGAGATTGATTTTCCATCTTTATTTCG
GAGAGCAGTGGAAGCGATGACGATGTGGGTACTACATTGCGCATTACCC
TACTTTGTGAGCGCATGGTTCGAAGCATGGATATGCGCATGCTGTGACTGC
CAAGATCCTCTACGCCGGACGCATCGTGGCCGGCATCACCGGCGCCACAG
GTGCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGGGGAAGATCGG
GCTCGCCACTTCGGGCTCATGAGCGCTTGTTCGGCGTGGGTATGGTGGC
AGGCCCCGTGGCCGGGGGACTGTTGGGCGCCATCTCCTTGATGCACCAT
TCCTTGCGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGGGCTGCTTC
CTAATGCAGGAGTCGCATAAGGGGCATCGGTGACGGGATCACGTTGTGTC
CCTGAAGCTCTCCTGTACCCAAACACAAAGGTGATGTCCCCAGCATCCCT
ATCCCAGCACTCTGGGGGACTCCTATTGAATTCCTCCTTGGGCTTGCTGC
CTTCTCTTCCCGTTCCAGAGATCCCAAAAGGTTAAGCACCTTTGGGTCA
GTGTTCAGAATTGTCACTGCCAGTTTGGGGTATCAGTGGCAAATTGAGA
CCCTTTTACCCAATCTTGCAACCACTCTGGTTCCCAAGTCTTATGGTTTTA
GATGGAGTAAAAAGGTTTATATGTCAAAAGTTCTTCTGTGTCTGGTTAT
TCGCTGCTTCTGGATGCCAGGATCATGGGGATAAGGGGAAAAACAATGGGT
TCTCTTATGCGTAGAGATGCAATCAGATGGGGAGAAAAAGAAATCTTAAT
CTTTCTGATCCATCTGACAGATATTCAGTACAGCCCTGAGGATGTGGGGA
AATAAATCTNGRAGAGTTKGTKGGCAGTTCCAAGGATTTGGGAATGACTA
AATCCCATTCCTGGKKWYTGACAAAGTTGSCTGTGTTGGAACCCAGAAA
GATCCATGCAAGTGGGTCACTCCCTGAAAGCATTGTGTTCTGCTGTCTGCT
AGCGGAGAGAAAGACACAGAGGGGAAAATTAAGTGTTTTATTGTTAATTA
TTGTACACTCTGAGGTTTCAAATAACCAATCTTTAACGAGAGCGGACCAC
TTGATTTGAGGGTGACCATCTCAGATGGGGACAACCTGTACCTGATCAGGC
AAACCTGGGGGAAATTTGCCTTTCTGCCACTCTTTTGGGTGGGATTTTCC
CTTTTGACCACCATTTTCTACATTCTAATCACCCATTGCAGCACTTCTCC
CCCTTTTTTTTGCCCCATTTTTCTCCTGCTCAGCACTTCTTAACAATATA
ATATAAATCAATATCATATCAATATGATTCTATGCCAATAGATTAATGGG
GATGAAAGACACATAAAAAACCCAAGTCCTCATTTTCATCTGCTTCCCATGG
GATGGGTGGGGAGGTGGCTGTCCCTGAGGCTGTAGGATGTGGGGTCAACC
CTTGTCTGTGTCTCAGGGACACAGCCTCAGCTTGGACCTGACCCCTACCA
CCCACAGCCACGGACGGACCCTCTCCCAGAGAAGGATGCATGGGAAAAA

CGGTCAGCCCATTTCGCCGCCAAGCTCTTCAGCAATATCACGGGTAGCCAA
CGCTATGTCCTGATAGCGGTCCGCCACACCCAGCCGGCCACAGTCGATGA
ATCCAGAAAAGCGGCCATTTTCC. CCATGATATTTCGGCAAGCAGGCATCG
CCATGGGTACGACGAGATCCTCGCCGTGGGCATGCGCGCCTTGAGCCT
GGCGAACAGTTCGGCTGGCGCGAGCCCCTGATGCTCTTCGTCCAGATCAT
CCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGA
TGTTTCGCTTGGTGGTTCGAATGGGCAGGTAGCCGGATCAAGCGTATGCAG
CCGCCGCATTGCATCAGCCATGATGGATACTTTCTCGGCAGGAGCAAGGT
GAGATGACAGGAGATCCTGCCCCGGCACTTCGCCCAATAGCAGCCAGTCC
CTTCCCCTTCAGTGACAACGTGAGCACAGCTGCGCAAGGAACGCCCGT
CGTGGCCAGCCACGATAGCCGCGCTGCCTCGTCTGCAAGTTCATTACGGG
CACCGGACAGGTTCGGTCTTGACAAAAAGAACCGGGCGCCCCCTGCGCTGAC
AGCCGGAACACGGCGGCATCAGAGCAGCCGATTGTCTGTTGTGCCAGTC
ATAGCCGAATAGCCTCTCCACCCAAGCGGCCGGAGAACCTGCGTGCAATC
CATCTTGTTCAATCATGCGAAACGATCCTCATCCTGTCTCTTGATCAGAT
CTGCGGCACGCTGTTGACGCTGTTAAGCGGGTCGCTGCAGGGTCGCTCGG
TATTCGAGGCCACACGCGTCACCTTAATATGCGAAGTGGACCTGGGACCG
CGCCGCCCCGACTGCATCTGCGTGTTTGAATTCGCCAATGACAAGACGCT
GGGCGGGGTTTGTGTATCATAGAACTAAAGACATGCAAATATATTTCTT
CCGGGGACACCGCCAGCAAACGCGAGCAACGGGCCACGGGGATGAAGCAG
CTGCGCCACTCCCTGAAGCTCCTGCAGTCCCTCGCGCCTCCGGGTGACAA
GATAGTGTACCTGTGCCCCGTCTGGTGTGTTGTCGCCAACGGACGCTCC
GCGTCAGCCGCGTGACCCGGCTCGTCCCGCAGAAGGTCTCCGGTAATATC
ACCGCAGTCGTGCGGATGCTCCAGAGCCTGTCCACGTATACGGTCCCCAT
TGAGCCTAGGACCCAGCGAGCCCGTCGCCCGCCGCGGCGCGCCCGCCGGG
GGTCTGCGAGCAGACCGAAAAGGTACACTCTGGGGCGCGCGACCCGCCC
GAGTCAGCGGCCCGCCAGTTACCACCCGCCGACCAAACCCCCGCTCCAC
GGAGGGCGGGGGGGTGTGTTAAGAGGATCGCGGCGCTCTTCTGCGTGCCCC
TGGCCACCAAGACCAAACCCCGAGCCGCTCCGAATGAGAGTGTTCGTT
CCTTCCCCCTCCCCCGCGTCAGACAAACCTAACCACCGCTTAAGCGGC
CCCCGCGAGGTCCGAAGACTCATTGATCGATCCGGAATTCTCATGTTT
GACAGCTTATCATCGATAAGCTTTAATGCGGTAGTTTATCACAGTTAAAT
TGCTAACGCAGTCAGGCACCGTGTATGAAATCTAACAATGCGCTCATCGT
CATCCTCGGCACCGTCACCCCTGGATGCTGTAGGCATAGGCTTGTTATGC
CGGTACTGCCGGGCTCTTGCGGGATATCGTCCATTCCGACAGCATCGCC
AGTCACTATGGCGTGCTGCTAGCGCTATATGCGTTGATGCAATTTCTATG
CGCACCCGTTCTCGGAGCACTGTCCGACCGCTTTGGCCGCCGCCAGTCC
TGCTCGCTTCGCTACTTGGAGCCACTATCGACTACGCGATCATGGCGACC
ACACCCGTCTGTGGATCTGCCTCGTTGGCTGCCGCAGTTCTTCAACCT
CCCGGCGCAGCTTTTCGTTCTCAATTTACGCATCCCTTTCCGGCATAACCAT
TTTATGACGGCGGCAGAGTCATAAAGCACCTCATTACCCTTGCCACCGCC
TCGCAGAACGGGCATTCCTGTCTGCTGCCAGTTCTGAATGGTACGGATAC
TCGCACCGAAAATGTCAGCCAGCTGCTTTTTGTTGACTTCCATTGTTCAAT
TCCACGGACAAAAACAGAGAAAGGAAACGACAGAGGCCAAAAAGCTCGCT
TTCAGCACCTGTGTTTTCTTTCTTTTCAGAGGGTATTTTAAATAAAAAAC
ATTAAGTTATGACGAAGAAGAACGGAAACGCCTTAAACCGGAAAAATTTTC
ATAAATAGCGAAAACCCGCGAGGTGCGCCGCCCGTAACAAGGCGGATCGC
CGGAAAGGACCCGCAAATGATAATAATTATCAATTGCATACTATCGACGG
CACTGCTGCCAGATAACACCACCGGGGAAACATTCCATCATGATGGCCGT
GCGGACATAGGAAGCCAGTTCATCCATCGCTTTCTTGTCTGCTGCCATTT
GCTTTGTGACATCCAGCGCCGCACATTCAGCAGCGTTTTTCAGCGCGTTT

Figure 6DM

GCACAAGGAA ATGCAAAGGG GCATCA TAG GGGACATGGC ACGGGGCATT
51 CTAGGGAGCA TTGCATGGGG ACATTGCAA GGAATGCAA AGGGACATTG
101 CATGGGGACA TTGCAAACAA ATTGAGTGGG AGATTGCACC GGGATGTTGC
151 ATGGGGACAT TGCATGGAAT GTCCCACCAA CCACCCTGCA GGGTGACACT
201 GGGACCATCC CCAGCTCTGA CCATCCCCC TTTGCTGCAG CACCACCCCA
251 GGTCCGCATC GTCCCCATCC CCATCTCCAA CGACCCCGAC ACCGTCCACC
301 TCATCTGCCA TGTTTGGGGC TTCTACCCAC CCGCAGTGAC CATCCAGTGG
351 CTGCACAACG GCCTCGTGGT GGCCTCAGGT GACACCAAAC TGCTGCCCAA
401 CGGGGGACTG GACCTACAGG ACACAGGTGG CCCTGAGGGC CAGCATTGCA
451 GCAGGGAGCA CTAAACATG TTCAGTGTGG CAATTCCAGC TTGGAGCAGC
501 CGCTGCAGGA GGATTGGAGT GAGTTTGGGG ATGGGGATGT GGCACCCACA
551 CCCCACAGTC CCCCACGGTT CATTGTGCC CACGCTCTCC CCACAGGTCC
601 CAATTTGTCC CCGGCGATGA TGGTGAAGGT GGCAGTGGCG GCCATGGCGC
651 TGACGTTGGG GTTGGTGGCA CTCAGCGCCG GGGTTTTCAG CTTCTGTCAG
701 CGGCCACGGG GTGAGGGATG GGGATGTGGT GCTGGGGACA TGTGTGACAC
751 CGAGGGTCTG GTGTCCAGTG TGGGGTGTAC CTCCTCATTC ATCATCTTCT
801 GTGTGGCAGC TCCTGGCGCT GGTCCCAGTA CCCCCTCCTG ATGCGGGTTC
851 TCACTCCAAT CCTGGTCCCC AAAATGATCC CGGTCCGAGT TCTGGTCCCC
901 ATCCCAGTCC TGGACCCCAT CCCAGTCCTG GTCCCCATTC TGGTCTTGGT
951 CCTGGTCTCG GTTCTGCTCC TGGTCCCTAT CCCTGACTCT GGTCCCGGTC
1001 CCCATCCCGA TGCCAGTCCC AGTCCTGGTC CCCATCCTGG TCCTGCTCCT
1051 TGGTTTGGGG ACCTCAATGA CTGGAAGTCC CATGTCCCAA CATGGGGACC
1101 CACAGTTTGG GGTGAGGGGC TCTACCCCC CAATAAAACC ATCTGCAGCC
1151 CCAACCTCGC TCCAATTCTT CGTTCCACG TTGGGTGGGT CGGGCTCCCA
1201 GTGCTCCCAG CCGTNTATGT CCCGTAAGCG TCGGCTCCAC TGCATAAAAA
1251 GAAAAAAAA AAA

Séquence Génomique TAP1
(Du début de l'exon 2 à l'extrémité 3')

GGC GAG ATG GCC GTG CCC TAC TAC ATG GGG CGA GCC AGC GAC TGG GTG GCC CGC GAG GAC AAG CTG GCA
G E M A V P Y Y M G R A S D W V A R E D K L A
GCC ATC CTG CCC ATG GTG CTG CTG GGC CTC AGC AG GTACTGGCATAGGGGGGACGCGGTGCGGGCAGGGGGCAGCGCG
I L P A M V L L G L S S
GGACCCCTGACACCCCACTGCCGTCACAG C GCT GTT ACT GAG CTG GTG TGT GAT GTG ACC TTC GTG GGG ACA
A V T E L V C D V T F V G T
CTG AGC CGC ACG CAA AGC CGC CTC CAG CGC CGC GTC TTC GCC GCC GTC CTG CGG CAG AGC ATC ACC GAG
L S R T Q S R L Q R R V F A A V L R Q S I T E
CTG CGC GCC GAT GGG GCC G GTGAGGGGCCACCGGGCTGGGAGGGGACACGGGGATAAGGGACAGGGGTGGCACTGACGGCGCTG
L R A D G A
TCACCCGGCAG GG GAT GTG GCC ATG CGG GTG ACG CGG GAT GCG GAG GAC GTG CGC GAG GCG CTG GGC AAG
G D V A M R V T R D A E D V R E A L G K
GCG CTG AGC CTC CTG CTG TGG TAT CTG GCA CGC GGC CTC TGC CTC TTT GCA ACC ATG GCC TGG CTG TCC
A L S L L L W Y L A R G L C L F A T M A W L S
CCG CGC ATG GCG CTG CTC ACC GCG CTG GCG CTG CCA CTG CTG CTG GCA CTG CCC AGG GCT GTG GGG CAC
P R M A L L T A L A L P L L L A L P R A V G H
TTC CGG CAG GTATGGGCTGCTGTCTGCACCTCCATGTGCCTTTGGTCCCCTCCATGTGCCTCTGGTCCCCTCCATGTGCCAGTGTG
F R Q
ACCACCATGTACTCATTGCCCTATCCATGTGCCCACTGTCCCCTCCATGTACCCACCATCCTCCTGCTGTGTGCTCCCCTCTGTGTGACCGGC
TGTCCCCTCCACGTGCCCCATGCCCTTCCATGCGTCCCACCATCCGTGCCATGTGCTCATTATTCCTATGTGTGACCAATTATCCCTTCCA
CATGGGTTCTGTTCTCTGCATGCCCCCACTGTCACTCCACATGCCACCGTCCCCTTTATGTCCCCTCCATCCCCTCCACGTGTTCTTTG
TTCCCCTCCATACATGCACTGTCCCCTCCCCAGCCCCCATTCCTTCCCACCCGCCCTGCAATGACACTGCTGTCCCCAG GCC CTG GCA
A L A
CCA CAG ATG CAG AAG GCG CAG GCC CGG GCC AGC GAG GTG GCA GTG GAG ACC TTC CAG GCC ATG GCC ACT
P Q M Q K A Q A R A S E V A V E T F Q A M A T
GTG CGC AGC TTT GCC AAT GAG GAT GGG GCA GCT GCA CAC TAC CGG CAG CGC CTG CAG CAG AGC CAC CGC
V R S F A N E D G A A A H Y R Q R L Q Q S H R
CTG GAG AAA AAG GAT GTG GCC CTC TAC ACT GCC TCT CTC TGG ACC AGT GGT GTATGGGATGGGGTGGCTCAAT
L E K K D V A L Y T A S L W T S G
AGCATGGGGACGTGATGGGATGGGGCTGGGGGATGTGGGGACATGATAGGATAGGACTGGGGGGCATGGGGACATGGTGGGATAGGGCTGG
GAGATGTGGAGACGTGATGTAATTGAGATGTCAGGAGATGGGGACAGAATGCCAACGGGCTGGAGGCCATAATGGTGTGGAGATGGCAGGT
CATGGGAATATGATGGCATGGGGACTGTGGGACATAGATTTGATGGCATGGGGACATCAGGATGTAGCAGGCACAACAGTTCAGGGGCTCT
GGGGCAGGAGGATGCAGTGACGTGGGAATGGGGCGGCATGGGGGCTCCAGGACACTGGGAACATGATGGCATGAGGGGACATAGCACAGAG
ATAGCACAGCTGTGGGACACTGGGACAGGGGGGACATTGACAGAACAGGAAGGTGACAGAGTGGTGCTGGGGACTCAGAGTCCCAGGGGGA
GGTGTCCCCTGGTGACCTCATGGCATCCTCAG TTC TCA GCC CTG GCC CTG AAG ATG GGG ATC CTC TAC TAT GGG
F S A L A L K M G I L Y Y G
GGG CAG CTG GTG GCC GCG GGG ACC GTC AGC ACT GGG GAC CTC GTC ACC TTC CTC CTC TAC CAG ATA CAG
G Q L V A A G T V S T G D L V T F L L Y Q I Q

2771422

TTC ACT GAT GTC CTG GAG GTGAGCCTGAGAGGATGCCCATATCCGCATGTCCCCATGTCTCCCTGCCACAGTCACAGTGTGTCA
F T D V L E
CAGTGTCCCTGTGTTCTACATCCTCCCTGCTGTGTCTCCAGGCCA TCGAGTGTGTCCCCATGCCCATCCATGTCCCTGTGTCCCCACG
TCTAACCCCTCTGTGACCCCTGCCACATCCCTGTGTACCTCCTCTGTCCCCATACCTACCCCGTCTGTCCCTATGCCAATCCCACTGTGT
CCTCTGCGTCCCCATGTACCATGCCACATGCCCTTGCCCTCCTGGCCAGTCACTACGCTGTCCCCAG GTC CTG CTC GAC TAC
V L L D Y
TTC CCC ACA CTG ATG AAG GCT GTG GGC TCT TCG GAA AAA ATC TTT GAG TTC CTG GAC CGG GAG CCA CAG
F P T L M K A V G S S E K I F E F L D R E P Q
GTC TCA CCC TCA GGG ACA ATG GCA CCC GCT GAC CTG CAG GGC CAC CTC CAG CTG GAG GAT GTC TGG TTC
V S P S G T M A P A D L Q G H L Q L E D V W F
TCC TAC CCT GGG CGC CAG GAA ACC CGT CCT CAA GTGGGCACAGACACAGCCAGGGGACACGGGGGTGTGGTGGGACA
S Y P G R Q E T R P Q
GCGTGACAGGTGTGGAGCACAGTGGGGTGATTGAGGGACATGGATGTGATGGACAGGGGTGTGAGGATATGAAACAAGGAGATACATGGAGG
GGGTGGTATGGGGACACTGGAGAGGGACATGAGATCATGGTATTGAGGGCGCGGGGACATGGCACATGGTGGGTTTGTGGCACTGGGACAT
GATGAGTGACACAGAGACATGGTGGGGAGGGCATGGGAATGTAGAGGCCGTGGTA
GGG GTA TCA CTG GAG CTG CGC CCC GGG GAG GTG CTG GCA CTG CTG GGA CCC CCG GGC GCA GGG AAG AGC
G V S L E L R P G E V L A L L G P P G A G K S
ACT CTG GTG GCC CTC GTG TCC CGC CTG CAC CAG CCC ACG GCC GGG CGC CTG CTG CTG GAT GGC CAC CCC
T L V A L V S R L H Q P T A G R L L L D G H P
CTC CCC GCC TAC CAG CAC TCC TAC CTG TGC CGC CAG GTGAGCAGCCACATGTCCCCATGGCTCCTGGTTGTCCCCCTG
L P A Y Q H S Y L C R Q
TGTTCTTGATATCAGCAGCCATCCTCATTGAGTCACCAGATATCTGGGTCCCCAGCCATCACCACACACCCCTGATGTCTCTGCCATATCA
CCACTGTGTCCCTGCACTGTCCCCGGCCAAGTCCCCAACCATCCTGTGTCCCCAACCATCCCCACCATGTCCCCAGATGTCCCTGACACAT
CCCCAGCCATCCCCACCGCCTCCACTGCCACGTTGCCATGTTCCTCCAGCTGTCCCCCCCCACTGCAG GTG GCC GTC GTC CCC CAG
V A V V P Q
GAG CCG CTG CTT TTT GCC CGC TCA CTC CAC GCC AAC ATT TCC TAT GGG TTG GGG GGC TGC AGC CGG GCA
E P L L F A R S L H A N I S Y G L G G C S R A
CAG GTG ACA GCG GCC GCC CGC CGG GTG GGC GCC CAC GAC TTC ATC ACT CGC CTG CCC CAA GGC TAC GAC
Q V T A A A R R V G A H D F I T R L P Q G Y D
ACA G GTAAGCTGTCCCTTTCTGTTCCGGGTCCCTCCATGGTCCCTCCAGCCTGACCCCGCTCGTCCCCGCAG AG GTG GGC
T E V G
GAG TTG GGA GGA CAG CTC TCC GGG GGG CAG CGG CAG GCG GTG GCC ATT GCC CGT GCA CTG CTG CGG GAC
E L G G Q L S G G Q R Q A V A I A R A L L R D
CCC CGC ATC CTC ATA CTC GAC GAG CAC ACC AGC GCC CTG GAC AAT GAG AGC CAG CAG CAG GTGGGATGTC
P R I L I L D E H T S A L D N E S Q Q Q
CCCCACGTCCCCGTGTCCCCACATCCCCCTGAGCCCTGTGTTCCCTCAGATTGCACGCCTAGGTCCCCATGGTCCCTGTTCTGGTGTCCC
CTTATCTCCACTCCTGGTGTCCCTCGGTCCCTGGCAGTGGCTGAGGAACATCCCCCTGAACCGTTTCTCCTCCACAG GTG GAG CAG
V E Q
GAG ATC CTC GCA GCC AAA GGG TCG GGG CGT GCA GTG CTG ATG GTG ACG GGG CGG GCA GCC CTG GCG GCG
E I L A A K G S G R A V L M V T G R A A L A A

Figure 7

27/31

2771422

CGG GCA CAA CGA GTG GTG GTG TTG GAG GGG GGA GAG GTG CGG CAG GAG GGA CCC CCC CAC GAG GTG GTG
R A Q R V V V L E G E V R Q E G P P Q E V V

CGC CCC GTC AGC CTT NTT GCG GGA CTG GGG ACA ACA AGG GAG CAC CGG GGG AGG GGG ACA GAG GGA TAG
R P V S L ? A G L G T T R E H R G R G T E G *

CGGGAGTTTTGGATGGGGAGGGCAGGGGGTGGGTGGGATGTGGGATGGGGACACTGCGCGTTGGGGACACTGAGGGTGGAGGTGGGGACAC
CGGGGCAGCAACAAGGGACCACAAGAGCTGTGCCGTGGGCACATGGATGCCGAGCCGGGCGCGCTGCCGTACCGCTGCTGTACGACACACA
ACGGCCACAGCATGGACTGCAGTGCCACTGAGTGCCACCAGGGCCGGGGGTGGGACACAGAAGTGGGAATAAGCCGCATGTTTGT

Figure 8

TAP2G

-213

CGCCATACATTNTGCGCCTGTCATGCACGGTGNTAATGGCCGACCTGGCCNTCATGTTGGCCCTGGCCANTTCTTCCCAGCACTGGCC
TTGGGCTGGGTGGG -107

-106

TCCTGGCGGAGGCCGGGCTGCGCCTCCTGGTGCTGGGGGGGGCCGGGAGCTGCTGGCCCCAGGGGACCCCGTGGGGCTGCAGTGCTGC
AGCATGGGCCCCGCC 0

t

+1 ATCTTCCTGACCCTACGGGGCTATGTAGGTCTGCCTGGAGCTCCCCCGGTGCTGCTGGCC ATG GCA ACG CCG TO
TGG CTG GTG CTG ACC CAC +93 5'UT

W L V L T H

M A T P S

+94

GGG ACA GCT GTG GTG GCA TTG CTC ACC TGG AGC CTC CTG GTC CCC ACT GTG GCC ACT GGG
GCA AAG GAG GCA AAG GCC TGG +174
G T A V V A L L T W S L L V P T V A T G
A K E A K A W

+175

GTG CCC CTG AGG CGG CTG CTG GCC CTC GCC TGG CCC GAG TGG CCG TTC CTT GGC TGT GCC
TTC CTC TTC CTC GCA TTG GCT +255
V P L R R L L A L A W P E W P F L G C A
F L F L A L A

+256

GCA CTG GGT GAG ACC TCA TTG CCC TAC TGC ACC GGG AGG GCT GTG GAT GTC CTC CGC CAG
GGG GAC GGC CTC GCC GCC TTC +336
A L G E T S L P Y C T G R A V D V L R Q
G D G L A A F

+337

ACC GCT GCT GTC GGC CTC ATG TGC CTG GCC TCT GCC AGC AG
GTAGGGACCCACATCCCTCCACAAACCCATCCACCTCTGGTGGTCTCT +429
T A A V G L M C L A S A S S

+430

GGTGGGTTTGGGGCTCTCTGTCCATATCTGGGGGTCTCTGATGGGTTCTGGGCACTCCACTGACCCTTTGTGATTCTCTGAAGGGTTCTG
GCTCTCCATGACCC +536

+537

CTGATGGGTTTTGGAGTGCAGCCCCCAATTCCTTCCAG C TCG CTG TTT GCC GGC TGC CGC GGT GGC CT
TTC ACC TTC ATC AGG TTC +624
F T F I R F S L F A G C R G G L

+625

CGC TTC GTC TTG CGC ACC CGC GAC CAG CTC TTC TCC AGC CTG GTG TAC CGG GAC CTC GCC
TTC TTC CAG AAC ACC ACA GCA +705
R F V L R T R D Q L F S S L V Y R D L A
F F Q N T T A

+706

G GTACAGACTGGGGGCACTTTTGTCCCTGTCCCCACACCATAACCCAGCTCACCCTACTCAACTCCACAG CT GAG
TTG GCC TCC CGG CTG ACC ACC +828
L A S R L T T A E

+829

GAT GTG ACG CTG GCG AGC AAC GTG TTG GCA CTC AAT ATC AAC GTC ATG CTG AGG AAC CTG
GGG CAG GTG CTG GGG CTC TGC +909
D V T L A S N V L A L N I N V M L R N L
G Q V L G L C

+910

GCC TTC ATG CTG GGG CTG TCC CCG CGC CTG ACA ATG CTG GCA CTG CTC GAA GTG CCG CTC
GCC GTC ACC GCA CGG AAA GTC +990
A F M L G L S P R L T M L A L L E V P L
A V T A R K V

+991 TAT GAC ACC CGG CAC CAG
GTGATAGCAGGGATGGGATGGTAGGGTTGGGGTGACAGGGATGGAGGCAATGGCAATGGGATGGGAACAGTGGGAGTGGGGAT +1091
Y D T R H Q

+1092
AGTGAGGTGGGGATTGTGGGGTCAGGGTGGCAGGGATGAGGGCAGCTGCAATGGGATGGGAACAGTGGGAATGGGGAGAGCAGGATGGGG
CATGGGTCCAACACA +1198

+1199
GCAAGGATGAGAGGATGGAGAAGAGTGGAGCAGGAATGGAAGTGGGATGGCGAGTACTTGGCCATCCCATGGGTGCTGACACCCACTGTCC
CCCCAG ATG CTG +1302

M L

+1303 CAG CGG GCC GTG CTG GAT GCA GCA GCC GAC ACC GGA GCG GCA GTG CAG GAG TCC ATC TCT
TCC ATT GAG ATG GTA CGG GTC +1383
Q R A V L D A A A D T G A A V Q E S I S
S I E T V R V

+1384 TTC AAT GGC GAG GAG GAG GAG GAG CAC CGC TAC AGC CAG GTG CTG GAC AGG ACC CTA CGG
CTG CGG GAC CAG CGG GAC ACA +1464
F N G E E E E E H R Y S Q V L D R T L R
L R D Q R D T

+1465 GAG AGG GCC ATT TTT CTC CTC ATC CAG CGG
GTGAGGCTGACACGAGGGGACACCCTGGTGTCTGGGTGGGATCGGGACATCCCCGCTGAGCCCCAT +1561
E R A I F L L I Q R

+1561 CCCCACAG GTG CTG CAG TTG GCT GTG CAG GCA CTG GTG CTA TAC TGT GGG CAC CAG CAG CTC
CGC GAA GGG ACC CTC ACT +1641
V L Q L A V Q A L V L Y C G H Q Q L
R E G D L T

+1642 GCC GGC AGC CTC GTC GCC TTC ATC CTC TAC CAG ACT AAA GCT GGC AGC TGC GTG CAG
GTGAGGTCAGGCAGTGCCTCTGCCACCG +1729
A G S L V A F I L Y Q T K A G S C V Q

+1730
GATCCCCATGACTGTGGCCACATCCCCGTGTCCCCACCCTGGGTGCTGTGCCTGGGGGTCACATCCCCATGTCCCTATCTGGGTGCTGTG
CATGCAG GCA CTG +1834

A L

+1835 GCG TAC TCC TAT GGT GAC CTT CTG AGC AAT GCA GTG GCC GCC TGC AAG GTC TTT GAT TAC
CTG GAC TGG GAG CGA CCT GTG +1915
A Y S Y G D L L S N A V A A C K V F D Y
L D W E R P V

+1916 GGT GCT GGT GGC ACC TAT GTG CCC ACC AGA CTG CGG GGC CAC ATC ACC TTC CAT CGG GTG
TCC TTC GCC TAT CCC ACT CGC +1996
G A G G T Y V P T R L R G H I T F H R V
S F A Y P T R

+1997 CCT GAG CGC CTC GTC CTG CAA GAT GTC ACC TTC GAG CTG CGC CCC AGT GAG GTG ACG GCG
TTG GCG GGG CTG AAT GGC AGC +2077
P E R L V L Q D V T F G L R P S E V T A
L A G L N G S

+2078 GGG AAG AGC ACC TGC GTG GCA CTG CTG GAG AGA TTC TAT GAA CCT GGG GCC GGG GAA GTG
CTG CTG GAC GGG GTG CCG CTG +2158
G K S T C V A L L E R F Y G P G A G E V
L L D G V P L

+2159 CGG GAC TAC GAG CAT CGC TAC CTG CAC CGC CAG
 GTGANGGGGTGGGGGAAATGTTAGCTGCACTGAACANTGCTGGGGCTGAACCTCTGCCCTGG +2254
 R D Y E H R Y L H R Q

+2255 GGGCAG GTG GCA CTG GTG GGG CAG GAA CCC GTG CTC TTC TCT GGC TCC ATT CGG GAT AAC
 ATT GCC TAC GGG ATG GAG GAC +2335
 V A L V G Q E P V L F S G S I R D N
 A Y G M E D

+2336 TGC GAA GAG GAG GAG ATC ATA GCA GCT GCA AGG GCT GCG GGT GCT TTG GGC TTC ATC TCT
 GCA CTG GAG CAA GGC TTT GGC +2416
 C E E E E I I A A A R A A G A L G F I S
 A L E Q G F G

+2417 ACT G GTGAGTGCTGGGGAGCAAGGGGGGACCCGGGTGTCTGACCCCACTCATCCCCACCCTCATCCTGCAG AC
 GTA GGG GAG AGA GGG GGG CAG +2511
 T
 G E R G G Q D

+2512 CTG TCA GCG GGG CAG AAG CAG CGC ATC GCC ATC GCC CGC GCT TTG GTG CGG CGT CCC ACC
 ATC CTT ATC CTC GAC GAA GCC +2592
 L S A G Q K Q R I A I A R A L V R R P T
 I L I L D E A

+2593 ACC AGT GCT CTG GAT GGG GAC AGC GAT GCA ATG
 GTGAGCACTGAGCAGTGGGTGGGGGGAGGGTCTG?CCCTGCAGTGCATGCTGATGGGCAGCTG +2688
 T S A L D G D S D A M

+2689 TGTGTCCTACAG CTA CAG CAG TGG GTG AGG AAC GGA GGG GAC CGG ACG GTG TTG TTT ATC ACC
 CAC CAA CCA CGG ATG CTG +2769
 L Q Q W V R N G G D R T V L F I T
 H Q P R M L

+2770 GAG AAG GCA GAC CGC ATT GTG GTG CTG GAG CAT GGC ACG GTG GCT GAG ATG GGG ACA CCC
 GCC GAG CTG AGG ACC CGC GGC +2850
 E K A D R I V V L E H G T V A E M G T P
 A E L R T R G

+2851 GGA CCC TAC AGC CGG CTG TTA CAG CAC TGA
 GAACCATGGAGCAGCTGGAGTGGCATGCGATGGGATATGGGGAGCAGTGACTGCCTTTGCTTCCAGC +2947
 G P Y S R L L Q H *

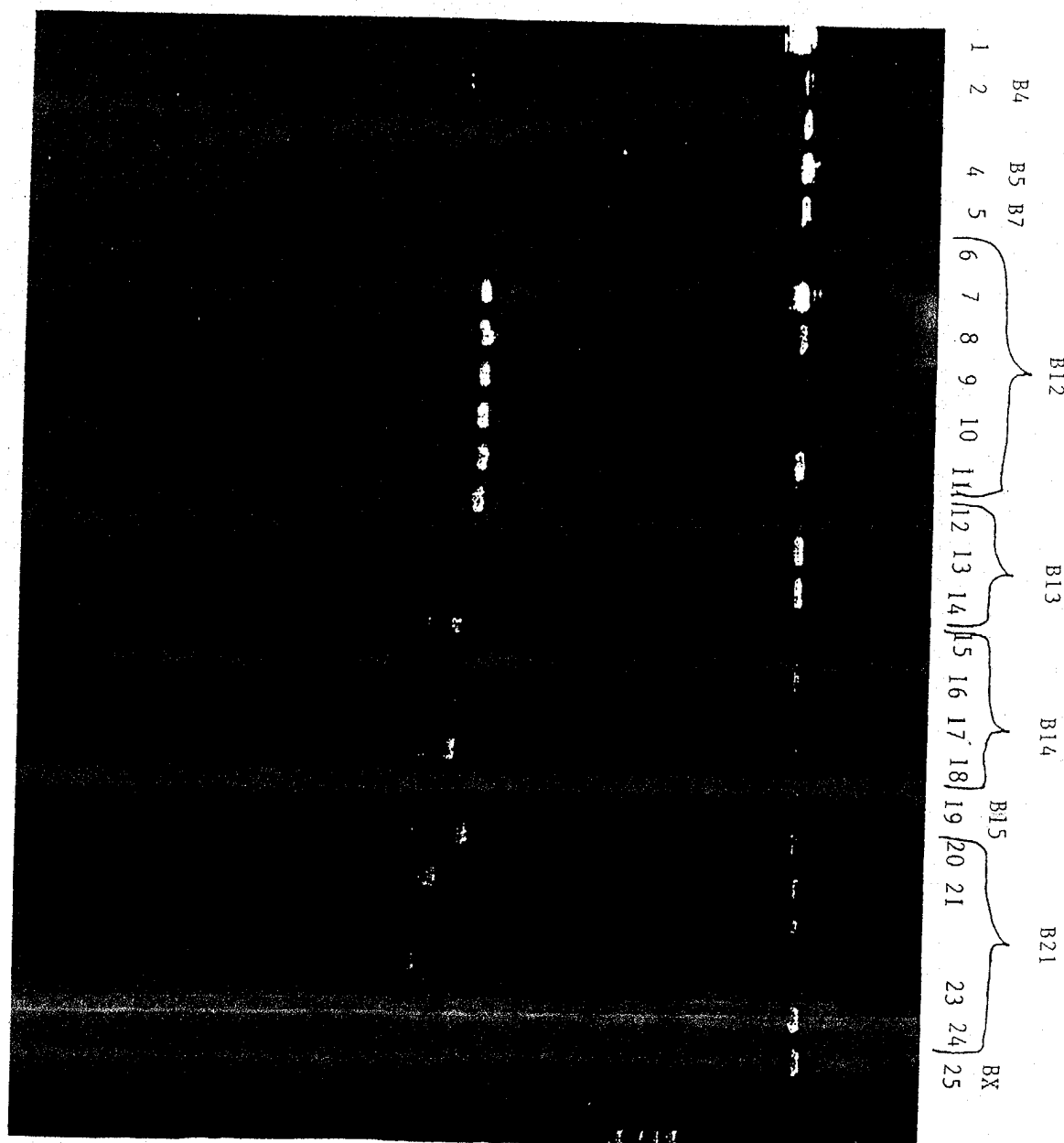
+2948 TGCAGGATGGGATGTTTTGGGATTTGTGTGGAATAAAGTGGAGATGCTTTGT
 +2999 3'UT

INTRON 2-3 : EF23(1B+1R)3R
 INTRON 3-4 : EF23 2(1)
 INTRON 4-5 : EF23 352H CON
 INTRON 5-6 : EF23224RS
 INTRON 6-7 : EF23(5B+5R)1R
 INTRON 8-9 : EF23277B CON
 INTRON 9-10: EF23 43RSR
 INTRON 10-11: EF23 43RSR
 INTRONS 1-2 ET 7-8 INEXISTANTS CHEZ LE POUET

2771422

31/31

Figure 9



REPUBLIQUE FRANÇAISE

2771422

INSTITUT NATIONAL
de la
PROPRIETE INDUSTRIELLE

RAPPORT DE RECHERCHE
PRELIMINAIRE
établi sur la base des dernières revendications
déposées avant le commencement de la recherche

N° d'enregistrement
national

FA 550336
FR 9714669

DOCUMENTS CONSIDERES COMME PERTINENTS			Revendications concernées de la demande examinée
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes		
X	GUILLEMOT F. ET AL.,: "Physical linkage of a guanine nucleotide-binding protein-related gene to the chicken major histocompatibility complex" PROC. NATL. ACAD. SCI. USA, vol. 86, - juin 1989 pages 4594-4598, XP002074404 le document en entier, esp. discussion ---	1	<div>DOMAINES TECHNIQUES RECHERCHES (Int.CL.6)</div> <div>C12Q</div>
X	GUILLEMOT F. ET AL.,: "A molecular map of the chicken major histocompatibility complex: the class II beta genes are closely linked to the class I genes and the nucleolar organizer" THE EMBO JOURNAL, vol. 7, no. 9, - 1988 pages 2775-2785, XP002074405 le document en entier, esp. figure 1 et page 2783, 2. colonne ---	1	
Y	MILLER M. M. ET AL.,: "Regions of homology shared by Rftp-Y and major histocompatibility B complex genes" IMMUNOGENETICS, vol. 39, - 1994 pages 71-73, XP002074406 * le document en entier *	1-10	
Y	WAKENELL P. S. ET AL.,: "Association between the Rfp-Y haplotype and the incidence of Marek's disease in chicken" IMMUNOGENETICS, vol. 44, - 1996 pages 242-245, XP002074407 * le document en entier *	1-10	
-/--			

1
EPO FORM 1503 (03.82 (P04C13))

Date d'achèvement de la recherche

13 août 1998

Examineur

Müller, F

CATEGORIE DES DOCUMENTS CITES

- X : particulièrement pertinent à lui seul
- Y : particulièrement pertinent en combinaison avec un
autre document de la même catégorie
- A : pertinent à l'encontre d'au moins une revendication
ou arrière-plan technologique général
- O : divulgation non-écrite
- P : document intermédiaire

- T : théorie ou principe à la base de l'invention
- E : document de brevet bénéficiant d'une date antérieure
à la date de dépôt et qui n'a été publié qu'à cette date
de dépôt ou qu'à une date postérieure.
- D : cité dans la demande
- L : cité pour d'autres raisons

& : membre de la même famille, document correspondant

REPUBLIQUE FRANÇAISE

2771422

INSTITUT NATIONAL

RAPPORT DE RECHERCHE

N° d'enregistrement
national

de la

PRELIMINAIRE

PROPRIETE INDUSTRIELLE

établi sur la base des dernières revendications
déposées avant le commencement de la recherche

FA 550336
FR 9714669

DOCUMENTS CONSIDERES COMME PERTINENTS		Revendications concernées de la demande examinée
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	
A	ZOOROB R. ET AL.,: "Chicken major histocompatibility complex class II B genes: analysis of interallelic and interlocus sequence variance" EUR. J. IMMUNOL., vol. 23, - 1993 pages 1139-1145, XP002074408 * le document en entier *	
A	VALLEJO R.L. ET AL.,: "Non-association between Rfp-Y major histocompatibility complex-like genes and susceptibility to Marek's disease virus induced tumours in 6.3x7.2 intercross chickens" ANIMAL GENETICS, vol. 28, - 5 octobre 1997 pages 331-337, XP002074409 * le document en entier *	
A	KAUFMAN J. ET AL.,: "Different features of the MHC class I heterodimer have evolved at different rates" J. IMMUNOLOGY, vol. 148, - 1 mars 1992 pages 1532-1546, XP002074410 * le document en entier *	

DOMAINES TECHNIQUES
RECHERCHES (Int.CL.6)

Date d'achèvement de la recherche

Examineur

13 août 1998

Müller, F

CATEGORIE DES DOCUMENTS CITES

X : particulièrement pertinent à lui seul
Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie
A : pertinent à l'encontre d'au moins une revendication ou arrière-plan technologique général
O : divulgation non-écrite
P : document intercalaire

T : théorie ou principe à la base de l'invention
E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure.
D : cité dans la demande
L : cite pour d'autres raisons

& : membre de la même famille, document correspondant